Appendix D
NATIONAL TOXICOLOGY PROGRAM'S REVISED MONOGRAPH ON FLUORIDE NEUROTOXICITY

National Toxicology Program
U.S. Department of Health and Human Services

DRAFT NTP MONOGRAPH ON THE

SYSTEMATIC REVIEW OF FLUORIDE EXPOSURE AND NEURODEVELOPMENTAL AND COGNITIVE HEALTH EFFECTS*

Revised September 16, 2020

*The September 6, 2019 draft monograph was peer reviewed by a committee convened by the National Academy of Sciences, Engineering, and Medicine (NASEM). This current draft incorporates changes in response to that review and is being submitted to the same NASEM committee for an additional round of peer review.

Office of Health Assessment and Translation
Division of the National Toxicology Program
National Institute of Environmental Health Sciences
National Institutes of Health

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

FOREWORD

The National Toxicology Program (NTP), established in 1978, is an interagency program within the Public Health Service of the U.S. Department of Health and Human Services. Its activities are executed through a partnership of the National Institute for Occupational Safety and Health (part of the Centers for Disease Control and Prevention), the Food and Drug Administration (primarily at the National Center for Toxicological Research), and the National Institute of Environmental Health Sciences (part of the National Institutes of Health), where the program is administratively located. NTP offers a unique venue for the testing, research, and analysis of agents of concern to identify toxic and biological effects, provide information that strengthens the science base, and inform decisions by health regulatory and research agencies to safeguard public health. NTP also works to develop and apply new and improved methods and approaches that advance toxicology and better assess health effects from environmental exposures.

NTP conducts literature-based evaluations to determine whether exposure to environmental substances (e.g., chemicals, physical agents, and mixtures) may be associated with adverse health effects. These evaluations result in hazard conclusions or characterize the extent of the evidence and are published in the NTP Monograph series, which began in 2011. NTP Monographs serve as an environmental health resource to provide information that can used to make informed decisions about whether exposure to a substance may be of concern for human health.

NTP conducts these health effects evaluations following pre-specified protocols that apply the general methods outlined in the "Handbook for Conducting a Literature-Based Health Assessment Using the OHAT Approach for Systematic Review and Evidence Integration."[1] The protocol describes project-specific procedures tailored to each systematic review in a process that facilitates evaluation and integration of scientific evidence from published human, experimental animal, and mechanistic studies.

The key feature of the systematic review approach is the application of a transparent framework to document the evaluation methods and the basis for scientific judgements. This process includes steps to comprehensively search for studies, select relevant evidence, assess individual study quality, rate confidence in bodies of evidence across studies, and then integrate evidence to develop conclusions for the specific research question. Draft monographs undergo external peer review prior to being finalized and published.

NTP Monographs are available free of charge on the NTP website and cataloged in PubMed, a free resource developed and maintained by the National Library of Medicine (part of the National Institutes of Health). Data for these evaluations are included in the Health Assessment and Workspace Collaborative.

For questions about the monographs, please email NTP or call 984-287-3211.

[1] OHAT is the abbreviation for Office of Health Assessment and Translation, which is within the Division of the National Toxicology Program at the National Institute of Environmental Health Sciences.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

TABLES:

FIGURES:

DATA FIGURES:

RISK-OF-BIAS FIGURES:

META-ANALYSES TABLES:

Attachment A

Attachment B

META-ANALYSES FIGURES:

Attachment A

ABSTRACT

Background: The overall objective of this evaluation was to undertake a systematic review of published literature to reach conclusions concerning the potential for exposure to fluoride to affect neurodevelopment and cognition. The review only addresses whether exposure to fluoride could present a potential hazard (i.e., has the potential to cause harm, at any exposure level, including exposures that are higher than those typically encountered from consuming fluoridated drinking water in the United States). Benefits of fluoride with respect to oral health are not addressed in this monograph.

Previous reviews of epidemiological studies, including a 2006 evaluation by the National Research Council (NRC), found support for an association between consumption of high levels of naturally occurring fluoride in drinking water and adverse neurological effects in humans and recommended further investigation (NRC 2006). Most of the evidence reviewed was from dental and skeletal fluorosis-endemic regions that have higher levels of naturally occurring fluoride than the fluoride concentrations historically added to water in community water fluoridation programs (0.8–1.2 mg/L). For community water systems that add fluoride, the Public Health Service now recommends a fluoride concentration of 0.7 mg/L.

NTP previously published a systematic review of the evidence from experimental animal studies of the effects of fluoride on learning and memory (NTP 2016). The systematic review found a low-to-moderate level of evidence that learning and memory deficits occur in non-human mammals exposed to fluoride. Studies in animals generally used fluoride drinking water concentrations that far exceeded the levels used in water fluoridation, and the lack of studies at lower fluoride concentrations was identified as a data gap. The evidence for effects on learning and memory was strongest (moderate) in animals exposed as adults, and evidence was weaker (low) in animals exposed during development. Since the publication of the NTP (2016) systematic review of the animal evidence, additional animal studies have been published, many examining the effects of perinatal exposures. In addition, the number of studies examining cognitive and neurobehavioral effects of fluoride in humans has grown considerably since the NRC (2006) review, including several recent prospective cohort studies evaluating prenatal fluoride exposures.

Objective: To conduct a systematic review of the human, experimental animal [extending (NTP 2016) report], and mechanistic literature to evaluate the evidence and develop hazard conclusions about whether fluoride exposure is associated with neurodevelopmental and cognitive effects.

Method: A systematic review protocol was developed and utilized following the Office of Health Assessment and Translation (OHAT) approach for conducting literature-based health assessments.

Results: The literature search and screening process identified 159 published human studies, 339 published experimental animal studies, and 60 in vitro/mechanistic studies relevant to the objective. Ninety-two of the 159 human studies evaluated the association between fluoride exposure and neurodevelopmental or cognitive effects, and the remaining human studies evaluated thyroid effects or other potential mechanistic data. The majority of the experimental animal studies were mechanistic studies, which were not assessed in the NTP (2016) report. Since the NTP (2016) systematic review (through April 2019), 35 experimental animal studies evaluating effects on learning and memory and/or motor activity and sensory effects of fluoride were identified.

Supported by a meta-analysis, the human body of evidence provides a consistent and robust pattern of findings that higher fluoride exposure (e.g., >1.5 mg/L in drinking water) is associated with adverse effects on neurocognitive development, including lower intelligence quotient (IQ) in children. There is a moderate level of evidence from cognitive neurodevelopmental studies in children based on five prospective cohort studies and 14 cross-sectional studies where exposure was identified as occurring prior to outcome. The evidence for cognitive effects in adults is limited, coming from two cross-sectional studies, and is inadequate to evaluate whether fluoride exposure in adults is associated with cognitive effects. The assessment of the new animal data focuses on evaluating a deficiency identified during the prior NTP (2016) review concerning the difficulty in distinguishing potential effects of fluoride on motor and sensory functions from effects specifically on learning and memory functions. Further examination of the animal data, including studies carried out at the NTP, has not resolved this issue. Because of this and other deficiencies related to overall study quality, the animal body of evidence is now considered inadequate to inform conclusions on whether fluoride exposure is associated with cognitive effects (including cognitive neurodevelopmental effects) in humans. While the animal data provide evidence of effects of fluoride on neurodevelopment, the human evidence base is primarily focused on cognitive neurodevelopmental effects, and these human data are the primary basis of conclusions.

Conclusions: Because the majority of available studies evaluated cognitive neurodevelopmental effects in children, the focus of the hazard conclusions is on cognitive neurodevelopmental effects, primarily IQ. When focusing on findings from studies with exposures in ranges typically found in drinking water in the United States (0.7 mg/L for optimally fluoridated community water systems)[2] that can be evaluated for dose response, effects on cognitive neurodevelopment are inconsistent, and therefore unclear. However, when considering all the evidence, including studies with exposures to fluoride levels higher than 1.5 mg/L in water, NTP concludes that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans*. This conclusion is based on a moderate level of evidence that shows a consistent and robust pattern of findings in human studies across several different populations demonstrating that higher fluoride exposure (e.g., >1.5 mg/L in drinking water) is associated with lower IQ and other cognitive effects in children. Limited and weaker evidence is considered to provide an inadequate level of evidence that fluoride is associated with cognitive effects in adults. The evidence from animal studies is inadequate to inform conclusions on cognitive effects, and the mechanisms underlying fluoride-associated cognitive neurodevelopmental effects are not well characterized.

[2]As of April 2020, 1.08% of persons living in the United States (~ 3.5 million people) were served by community water systems (CWS) containing ≥ 1.1 mg/L naturally occurring fluoride. CWS supplying water with ≥ 1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population (~ 1.9 million people), and systems supplying water with ≥ 2 mg/L naturally occurring fluoride served 0.31% of the U.S. population (~1 million people) (https://www.cdc.gov/fluoridation/data-tools/reporting-system.html).

INTRODUCTION

The NTP's Office of Health Assessment and Translation (OHAT) conducted a systematic review to evaluate the evidence that exposure to fluoride is associated with neurodevelopmental or cognitive effects. This review was initiated in response to a nomination from the Fluoride Action Network. There are numerous human and animal studies reporting neurodevelopmental and cognitive health effects of exposure to excess fluoride. As noted by the National Research Council (NRC) in their 2006 report, although the studies lacked sufficient detail to fully assess their quality and relevance to the U.S. populations, the consistency of the results suggesting that fluoride may be neurotoxic warrants additional research (NRC 2006).

Fluoride salts are added to community water systems and dental products in the United States (e.g., toothpaste, mouth rinses, and supplements) for the prevention of dental caries. Approximately 67% of the U.S. population receives fluoridated water through a community drinking water system (CDC 2013). In other countries fluoride supplementation has been achieved by fluoridating food products such as salt, or milk. Fluoride supplementation has been recommended to prevent bone fractures (Jones *et al.* 2005). Fluoride also can occur naturally in drinking water. Other sources of human exposure include other foods and beverages, industrial emissions, pharmaceuticals, and pesticides (e.g., cryolite, sulfuryl fluoride). Soil ingestion is another source of fluoride exposure in young children (US EPA 2010).

The U.S. Public Health Service (PHS) first recommended communities add fluoride to drinking water in 1962. PHS guidance is advisory, not regulatory, which means that while PHS recommends community water fluoridation as a public health intervention, the decision to fluoridate water systems is made by state and local governments.[3] For community water systems that add fluoride, PHS now recommends a fluoride concentration of 0.7 milligrams/liter (mg/L) (equal to 0.7 parts per million [ppm]). Under the Safe Drinking Water Act, the U.S. Environmental Protection Agency (EPA) sets maximum exposure level standards for drinking water quality. The current enforceable drinking water standard for fluoride, or the maximum contaminant level goal (MCLG, a concentration at which no adverse health effects are expected), is 4.0 mg/L. This is the maximum amount of fluoride contamination (naturally occurring not from water fluoridation) that is allowed in water from public water systems; it is set to protect against increased risk of skeletal fluorosis, a condition characterized by pain and tenderness of the major joints. EPA also has a non-enforceable secondary drinking water standard of 2.0 mg/L, which is recommended to protect children against the tooth discoloration and/or pitting that can be caused by severe dental fluorosis during the formative period prior to eruption of the teeth. Although the secondary standard is not enforceable, EPA does require that public water systems notify the public if the average levels exceed it (NRC 2006). As of April 2020, 1.08% of persons living in the United States (~ 3.5 million people) were served by community water systems (CWS) containing ≥ 1.1 mg/L naturally occurring fluoride. CWS supplying water with ≥ 1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population (~ 1.9 million people), and systems supplying water with ≥ 2 mg/L naturally occurring fluoride served 0.31% of the U.S. population (~1 million people) (https://www.cdc.gov/fluoridation/data-tools/reporting-system.html).

Controversy over community water fluoridation stems from concerns about the potential harmful effects of fluoride and the ethics of water fluoridation. Commonly cited health concerns related to

[3] For many years, most fluoridated community water systems used fluoride concentrations ranging from 0.8 to 1.2 mg/L (US DHHS 2015).

fluoride are bone fractures and skeletal fluorosis, lower intelligence quotient (IQ) and other neurological effects, cancer, and endocrine disruption. Effects on neurological function, endocrine function (e.g., thyroid, parathyroid, pineal), metabolic function (e.g., glucose metabolism), and carcinogenicity were assessed in the 2006 NRC report Fluoride in Drinking Water: A Scientific Review of EPA's Standards (NRC 2006). The NRC review considered adverse effects of water fluoride, focusing on a range of concentrations (2–4 mg/L) above the current 0.7-mg/L recommendation for community water fluoridation (NRC 2006). The NRC report concluded that the current MCLG should be lowered to protect against severe enamel fluorosis and to reduce the risk of bone fractures associated with skeletal fluorosis (NRC 2006). Other than severe fluorosis, the NRC did not find sufficient evidence of negative health effects at fluoride levels below 4.0 mg/L; however, the NRC concluded that the consistency of the results of IQ deficits in children exposed to fluoride at 2.5 to 4 mg/L in drinking water from a few epidemiological studies of Chinese populations appeared significant enough to warrant additional research on the effects of fluoride on intelligence. The NRC report noted several challenges to evaluating the literature, citing deficiencies in reporting quality, lack of consideration of all sources of fluoride exposure, incomplete consideration of potential confounding, selection of inappropriate control subject populations in epidemiological studies, absence of demonstrated clinical significance of reported endocrine effects, and incomplete understanding of the biological relationship between histological, biochemical, and molecular alterations with behavioral effects (NRC 2006).

In 2016, NTP conducted a systematic review of the evidence from experimental animal studies on the potential effects of fluoride exposure on learning and memory (NTP 2016). The NTP (2016) systematic review found a low-to-moderate level of evidence that learning and memory deficits occur in experimental animals exposed to fluoride. Based on the findings in NTP (2016), NTP decided to conduct additional animal studies before carrying out a full systematic review to incorporate human, animal, and potentially relevant mechanistic evidence in order to reach hazard identification conclusions for fluoride and learning and memory effects. As the NTP (2016) report on the experimental animal evidence focused on learning and memory and developed confidence ratings for bodies of evidence by life stage of exposure (i.e., exposure during development or adulthood), this report also evaluates two different age groups in humans (i.e., children and adults) with a focus on cognitive neurodevelopmental effects in children and cognitive effects in adults in order to address potential differences in the health impact based on timeframe of exposure (i.e., during development or during adulthood). This evaluation has been conducted separately from the 2016 experimental animal assessment, but like the 2016 assessment, it has assessed mainly learning and memory effects in experimental animal studies to determine whether the findings inform the assessment of cognitive neurodevelopmental effects in children and cognitive effects in adults. The September 6, 2019 draft of this monograph was reviewed by a committee convened by the National Academy of Sciences, Engineering, and Medicine (NASEM). The current document incorporates changes in response to that review.

OBJECTIVE AND SPECIFIC AIMS

Objective

The overall objective of this evaluation is to undertake a systematic review to develop NTP hazard identification conclusions on the association between exposure to fluoride and neurodevelopmental and cognitive effects based on integrating levels of evidence from human and non-human animal studies with consideration of the degree of support from mechanistic data.

Specific Aims

- Identify literature that assessed neurodevelopmental and cognitive health effects, especially outcomes related to learning, memory, and intelligence, following exposure to fluoride in human, animal, and relevant in vitro/mechanistic studies.
- Extract data on potential neurodevelopmental and cognitive health effects from relevant studies.
- Assess the internal validity (risk of bias) of individual studies using pre-defined criteria.
- Assess effects on thyroid function to help evaluate potential mechanisms of impaired neurological function.
- Summarize the extent and types of health effects evidence available.
- Describe limitations of the systematic review, limitations of the evidence base, identify areas of uncertainty, as well as data gaps and research needs for neurodevelopmental and cognitive health effects of fluoride.

Dependent on the extent and nature of the available evidence:

- Synthesize the evidence using a narrative approach or meta-analysis (if appropriate) considering limitations on data integration such as study design heterogeneity.
- Rate confidence in the body of evidence for human and animal studies separately according to one of four statements: High, Moderate, Low, or Very Low/No Evidence Available.
- Translate confidence ratings into level of evidence of health effects for human and animal studies separately according to one of four statements: High, Moderate, Low, or Inadequate.
- Combine the level of evidence ratings for human and animal data to reach one of five possible hazard identification conclusions: Known, Presumed, Suspected, Not classifiable, or Not identified to be a hazard to humans.

METHODS

Problem Formulation and Protocol Development

The research question and specific aims stated above were developed and refined through a series of problem formulation steps including:

(1) receipt of nomination from the public in June 2015 to conduct analyses of fluoride and developmental neurobehavioral toxicity;

(2) analysis of the extent of evidence available and the merit of pursuing systematic reviews, given factors such as the extent of new research published since previous evaluations and whether these new reports address or correct the deficiencies noted in the literature (NRC 2006, OEHHA 2011, SCHER 2011);

(3) request for information in a Federal Register notice (dated October 7, 2015);

(4) consideration of comments providing a list of studies to review through Federal Register notice and public comment period from October 7, 2015 to November 6, 2015;

(5) release of draft concept titled *Proposed NTP Evaluation on Fluoride Exposure and Potential for Developmental Neurobehavioral Effects* in November 2015;

(6) presentation of draft concept at the NTP Board of Scientific Counselors (BSC) meeting on December 1−2, 2015;

(7) consideration of comments on NTP's draft concept from the NTP BSC meeting in December 2015; and

(8) consideration of input on the draft protocol from review by technical advisors.

NTP published a systematic review of the animal evidence on the effects of fluoride on learning and memory (NTP 2016). NTP has conducted additional studies in animals to assess the effect of fluoride exposure on learning and memory. The results from this experimental animal work were published (McPherson *et al.* 2018) and are incorporated into the current review, which considers the epidemiological, animal, and mechanistic evidence in its conclusions. The protocol used to conduct this systematic review was posted in June 2017 with updates posted in May 2019 and September 2020 (https://ntp.niehs.nih.gov/go/785076).[4] A brief summary of the methods is presented below.

NASEM Review

The September 6, 2019 draft of this monograph was peer reviewed by a committee convened by NASEM. The current draft reflects clarifications and changes in response to that review (NASEM 2020), including the addition of meta-analyses of the IQ studies in children.

PECO Statements

PECO (**P**opulation, **E**xposure, **C**omparators and **O**utcomes) statements were developed as an aid to identify search terms and inclusion/exclusion criteria as appropriate for addressing the overall research question (effects on neurodevelopmental or cognitive function and thyroid associated with fluoride exposure) for the systematic review (Higgins and Green 2011).The PECO statements are listed below for human, animal, and in vitro/mechanistic studies (see **Table 1**, **Table 2**, and **Table 3**).

Using the PECO statements, the evaluation searched for evidence of neurodevelopmental or cognitive function, and thyroid effects associated with fluoride exposure from human studies, controlled exposure animal studies, and mechanistic/in vitro studies. Mechanistic data can come from a wide variety of studies that are not intended to identify a disease phenotype. This source of experimental data includes in vitro and in vivo laboratory studies directed at cellular, biochemical, and molecular mechanisms that attempt to explain how a substance produces particular adverse health effects. The mechanistic data were first organized by general categories (e.g., biochemical effects in the brain and neurons, neurotransmitters, oxidative stress, etc.) to evaluate the information available. Categories focused on were those with more robust data at levels of fluoride more relevant to human exposure. The intent was not to develop a mechanism for fluoride induction of learning and memory effects, but to evaluate whether a plausible series of mechanistic events exists to support effects observed in the low-dose

[4] NTP conducts systematic reviews following pre-specified protocols that describe the review procedures selected and applied from the general methods outlined in the OHAT Handbook for Conducting a Literature-Based Health Assessment (http://ntp.niehs.nih.gov/go/38673). The protocol describes project-specific procedures tailored to each systematic review.

region (below approximate drinking water equivalent concentrations of 20 ppm) that may strengthen the hazard conclusion.

Table 1. Human PECO (Population, Exposure, Comparator and Outcome) Statement	
PECO Element	**Evidence**
Population	Humans without restriction as to age or sex, geographic location, or life stage at exposure or outcome assessment
Exposure	Exposure to fluoride based on administered dose or concentration, biomonitoring data (e.g., urine, blood, other specimens), environmental measures (e.g., air, water levels), or job title or residence. Relevant forms are those used as additives for water fluoridation: • Fluorosilicic acid (also called hydrofluorosilicate; CASRN 16961-83-4) • Sodium hexafluorosilicate (also called disodium hexafluorosilicate or sodium fluorosilicate; CASRN 16893-85-9) • Sodium fluoride (CASRN 7681-49-4) • Other forms of fluoride that readily dissociate into free fluoride ions (e.g., potassium fluoride, calcium fluoride, ammonium fluoride)
Comparators	Comparable populations not exposed to fluoride or exposed to lower levels of fluoride (e.g., exposure below detection levels)
Outcomes	Neurodevelopmental outcomes including learning, memory, intelligence, other forms of cognitive behavior, other neurological outcomes (e.g., anxiety, aggression, motor activity), and biochemical changes in the brain or nervous system tissue; or measures of thyroid function, biochemical changes, or thyroid tissue

Table 2. Animal PECO Statement	
PECO Element	**Evidence**
Population	Non-human mammalian animal species (whole organism)
Exposure	Exposure to fluoride based on administered dose or concentration, and biomonitoring data (e.g., urine, blood, other specimens). Relevant forms are those used as additives for water fluoridation: • Fluorosilicic acid (also called hydrofluorosilicate; CASRN 16961-83-4) • Sodium hexafluorosilicate (also called disodium hexafluorosilicate or sodium fluorosilicate; CASRN 16893-85-9) • Sodium fluoride (CASRN 7681-49-4) • Other forms of fluoride that readily dissociate into free fluoride ions (e.g., potassium fluoride, calcium fluoride, ammonium fluoride)
Comparators	Comparable animals that were untreated or exposed to vehicle-only treatment
Outcomes	Neurodevelopmental outcomes including learning, memory, intelligence, other forms of cognitive behavior, other neurological outcomes (e.g., anxiety, aggression, motor activity), and biochemical changes in the brain or nervous system tissue; or measures of thyroid function, biochemical changes, or thyroid tissue

Table 3. In Vitro/Mechanistic PECO Statement	
PECO Element	**Evidence**
Population	Human or animal cells, tissues, or biochemical reactions (e.g., ligand binding assays)
Exposure	Exposure to fluoride based on administered dose or concentration. Relevant forms are those used as additives for water fluoridation: • Fluorosilicic acid (also called hydrofluorosilicate; CASRN 16961-83-4) • Sodium hexafluorosilicate (also called disodium hexafluorosilicate or sodium fluorosilicate; CASRN 16893-85-9) • Sodium fluoride (CASRN 7681-49-4) • Other forms of fluoride that readily dissociate into free fluoride ions (e.g., potassium fluoride, calcium fluoride, ammonium fluoride)
Comparators	Comparable cells or tissues that were untreated or exposed to vehicle-only treatment
Outcomes	Endpoints related to neurological and thyroid function, including neuronal electrophysiology; mRNA, gene, or protein expression; cell proliferation or death in brain or thyroid tissue/cells; neuronal signaling; synaptogenesis, etc.

Literature Search

Main Literature Search

Search terms were developed to identify all relevant published evidence on developmental neurobehavioral toxicity or thyroid-related health effects potentially associated with exposure to fluoride by reviewing Medical Subject Headings for relevant and appropriate neurobehavioral and thyroid-related terms, and by extracting key neurological and thyroid-related health effects and

developmental neurobehavioral terminology from reviews and a sample of relevant primary data studies. A combination of relevant subject headings and keywords were subsequently identified. A test set of relevant studies was used to ensure the search terms retrieve 100% of the test set. Six electronic databases were searched (see **Main Literature Database Search**) using a search strategy tailored for each database (specific search terms used for the PubMed search are presented in **Appendix 1**; the search strategy for other databases are available in the protocol (https://ntp.niehs.nih.gov/go/785076). A search of PubChem indicated that sodium fluoride was not found in either the Tox21 or ToxCast databases; therefore, these databases were not included in the search. No language restrictions or publication year limits were imposed. These six databases were searched in December 2016 and the search was regularly updated during the review process through April 1, 2019.

Evaluations must include cut-off dates for the literature search to enable synthesis and development of conclusions. Following the NASEM committee peer review in November 2019 (NASEM 2020), an additional search was conducted on May 1, 2020, where only primary human epidemiology studies were prioritized during screening. The review of the 2020 search results focused on the human studies because they formed the basis of the conclusions in the September 6, 2019 draft. A supplemental literature search of Chinese-language databases (described below) was also conducted.

Publications identified in these searches are categorized as "references identified through database searches" in **Figure 4.** Studies identified from other sources or manual review that might impact conclusions were considered under "references identified through other sources" in **Figure 4**. Literature searches for this systematic review were conducted independently from the literature search conducted for NTP (2016). The current literature search strategy was based on the search terms used for NTP (2016) and refined for the current evaluation, including the addition of search terms to identify human studies. Although the review process identified studies prior to 2015, the current assessment did not evaluate the studies published prior to 2015 and relied on the NTP (2016) assessment. The focus of the literature searches for this systematic review was to identify and evaluate relevant animal studies that were published since completion of the literature searches for the NTP (2016) assessment in addition to the human and mechanistic data that were not previously evaluated.

Supplemental Chinese Database Literature Search

Following NASEM committee peer review in November 2019 (NASEM 2020), additional searches were developed for non-English-language databases to systematically search for studies that were previously identified from other resources (e.g., Chinese-language studies from the Fluoride Action Network website). Non-English-language databases with the greatest potential to contain relevant non-English publications that were not previously identified through database searches were selected. Multiple non-English language databases were explored before finding two databases (CNKI and Wanfang) that covered studies previously identified from other sources. These two Chinese electronic databases were searched in May 2020 with no language restrictions or publication year limits. Search terms from the main literature search were refined to focus on human epidemiology studies. The CNKI and Wanfang databases have character limits in the search strings; therefore, key terms were prioritized using text analytics to identify the most prevalent terms from neurodevelopmental or cognitive human epidemiology studies previously identified as relevant. Search strings were designed to capture known relevant studies that were previously identified from searching other resources without identifying large numbers of non-relevant studies [the search strategy for both databases are available in the protocol (https://ntp.niehs.nih.gov/go/785076)]. Publications retrieved were compared to publications retrieved from the main literature search and duplicates were removed. The remaining relevant publications are categorized as "references identified through database searches" in **Figure 4**. New animal and

mechanistic references retrieved were scanned for evidence that might extend the information currently in the September 6, 2019 draft. Although additional studies were identified, data that would materially advance the animal and mechanistic findings were not identified; therefore, these studies were not extracted nor were they added to the draft. Newly-retrieved human references were reviewed to identify studies that might impact conclusions with priority given to identifying and translating null studies that may have been missed using previous approaches. Null studies that were identified were translated and included.

Databases Searched

Main Literature Database Search

- BIOSIS (Thomson Reuters)
- EMBASE
- PsycINFO (APA PsycNet)
- PubMed (NLM)
- Scopus (Elsevier)
- Web of Science (Thomson Reuters, Web of Science indexes the journal Fluoride)

Supplemental Chinese Database Literature Search

- CNKI
- Wanfang

Searching Other Resources

The reference lists of all included studies; relevant reviews, editorials and commentaries; and the Fluoride Action Network website (http://fluoridealert.org/) were manually searched for additional relevant publications. Following NASEM committee peer review in November 2019 (NASEM 2020), the Fluoride Action Network website was again searched for relevant references and contacted to identify null or no effect studies.

Unpublished Data

Unpublished data were eligible for inclusion provided the owner of the data was willing to have the data made public and peer reviewed (see protocol for more details https://ntp.niehs.nih.gov/go/785076). No unpublished data were identified during the literature search.

Study Selection

Evidence Selection Criteria

In order to be eligible for inclusion, studies had to satisfy eligibility criteria that reflect the PECO statement in **Table 1**, **Table 2**, and **Table 3**. The following additional exclusion criteria were applied (see protocol for additional details; https://ntp.niehs.nih.gov/go/785076):

(1) Case studies and case reports.

(2) Articles without original data (e.g., reviews, editorials, or commentaries). Reference lists from these materials, however, were reviewed to identify potentially relevant studies not identified from the database searches. New studies identified were assessed for eligibility for inclusion.

(3) Conference abstracts or reports.

Screening Process

References retrieved from the literature search were independently screened by two trained screeners at the title and abstract level to determine whether a reference met the evidence selection criteria. Screening procedures following the evidence selection criteria in the protocol were pilot-tested with experienced contract staff overseen by NTP. For citations with no abstract or non-English abstracts, articles were screened based on title relevance (title would need to indicate clear relevance); number of pages (articles ≤ 2 pages were assumed to be conference reports, editorials, or letters unlikely to contain original data); and/or PubMed Medical Subject Headings (MeSH). Using this approach, literature was manually screened for relevance and eligibility against the evidence selection criteria using a structured form in SWIFT-Active Screener. While the human screeners review studies, SWIFT-Active Screener aids in the process by employing a machine-learning software program used to priority-rank studies for screening (Howard *et al.* 2020). SWIFT-Active Screener also refines a statistical model that continually ranks the remaining studies according to their likelihood for inclusion. In addition, SWIFT-Active Screener employs active learning to continually incorporate user feedback during title and abstract screening to predict the total number of included studies, thus providing a statistical basis for a decision about when to stop screening (Miller *et al.* 2016). Title and abstract screening was stopped once the statistical algorithm in SWIFT-Active Screener estimated that 98% of the predicted number of relevant studies were identified.

Studies that were not excluded during the title and abstract screening were further screened for inclusion with a full-text review by two independent reviewers using DistillerSR® by Evidence Partners, a web-based, systematic-review software program with structured forms and procedures to ensure standardization of the process. Screening conflicts were resolved through discussion and consultation with technical advisor(s), if necessary. During full-text review, studies that were considered relevant were tagged to the appropriate evidence streams (i.e., human, animal, and/or in vitro). Studies tagged to human or animal evidence streams were also categorized by outcome as primary neurodevelopmental or cognitive outcomes (learning, memory, and intelligence); secondary neurobehavioral outcomes (anxiety, aggression, motor activity, or biochemical); or related to thyroid effects. In vitro data were tagged as being related to neurological effects or thyroid effects. Translation assistance was sought to assess the relevance of non-English studies. Following full-text review, the remaining studies were "included" and used for the evaluation.

Screening of the May 2020 Literature Search Update

Following the NASEM committee peer review in November 2019 (NASEM 2020), an additional search was conducted on May 1, 2020, where only primary human epidemiology studies were identified. The study screening and selection process was focused on the human studies with primary outcomes for the evaluation because they form the basis of the conclusions. Animal in vivo, human secondary outcome-only, and human and animal mechanistic references were identified as part of the screening process. These studies were then scanned for evidence that might extend the information in the September 6, 2019 draft. Studies from the May 2020 literature search update will be listed in an appendix; however, data from the studies were not extracted unless it was believed they would materially advance the findings.

Data Extraction

Data were collected (i.e., extracted) from included studies by one member of the evaluation team and checked by a second member of the team for completeness and accuracy. Any discrepancies in data extraction were resolved by discussion or consultation with a third member of the evaluation team.

Data extraction was completed using the Health Assessment Workspace Collaborative (HAWC), an open source and freely available web-based interface application.[5] Data extraction elements are listed separately for human, animal, and in vitro studies in the protocol (https://ntp.niehs.nih.gov/go/785076). Data for primary and secondary outcomes as well as thyroid hormone level data were extracted from human studies. Studies evaluating only goiters or thyroid size were not extracted. All primary outcomes and functional neurological secondary outcomes (e.g., motor activity) were extracted from animal studies identified since the NTP (2016) report. For animal mechanistic data, studies were tiered based on exposure dose (with preference given to fluoride drinking water equivalent exposures, which were calculated using the method described in the NTP (2016) report) of 20 ppm or less as deemed most relevant to exposures in humans), exposure duration or relevant time window (i.e., developmental), exposure route (with preference given to oral exposures over injection exposures), and commonality of mechanism (e.g., inflammation, oxidative stress, changes in neurotransmitters, and histopathological changes were considered pockets of mechanistic data). Data were not extracted from in vitro studies; however, these studies were evaluated for biological plausibility of the human and animal results. Thyroid data were also reviewed but not extracted. The data extraction results for included studies are publicly available and can be downloaded in Excel format through HAWC (https://hawcproject.org/assessment/405/). Methods for transforming and standardizing dose levels and results from behavioral tests in experimental animals are detailed in the protocol (https://ntp.niehs.nih.gov/go/785076).

In 2016, NTP conducted a systematic review of the evidence from experimental animal studies on the potential effects of fluoride exposure on learning and memory (NTP 2016). The literature searches for the current assessment identified and evaluated relevant animal studies published since the 2016 assessment and also included human and mechanistic data that were not previously evaluated. Although literature search activities for the current assessment identified experimental animal studies prior to 2015, the current assessment did not re-evaluate studies published prior to 2015, but relied on the NTP (2016) assessment.

Quality Assessment of Individual Studies

Risk of bias was assessed for individual studies using a tool developed by OHAT that outlines a parallel approach to evaluating risk of bias from human, animal, and mechanistic studies to facilitate consideration of risk of bias across evidence streams with common terms and categories. The risk-of-bias tool is comprised of a common set of 11 questions that are answered based on the specific details of individual studies to develop risk-of-bias ratings for each question. Study design determines the subset of questions used to assess risk of bias for an individual study (see **Table 4**).

Assessors were trained with an initial pilot phase undertaken to improve clarity of rating criteria and to improve consistency among assessors. Studies were independently evaluated by two trained assessors who answered all applicable risk-of-bias questions with one of four options in **Table 5** following pre-specified criteria detailed in the protocol (https://ntp.niehs.nih.gov/go/785076). The criteria describe aspects of study design, conduct, and reporting required to reach risk-of-bias ratings for each question

[5] HAWC (Health Assessment Workspace Collaborative): A Modular Web-based Interface to Facilitate Development of Human Health Assessments of Chemicals (https://hawcproject.org/portal/).

and specify factors that can distinguish among ratings (e.g., what separates "definitely low" from "probably low" risk of bias).

Key Risk-of-bias Questions

In the OHAT approach, some risk-of-bias questions or elements are considered potentially more important when assessing studies because there is more empirical evidence that these areas of bias have a greater impact on estimates of the effect size or because these issues are generally considered to have a greater effect on the credibility of study results in environmental health studies (Rooney *et al.* 2016). There were three Key Questions for observational human studies: confounding, exposure characterization, and outcome assessment. There were also three Key Questions for experimental animal studies: randomization, exposure characterization, and outcome assessment. In addition, for animal developmental studies, failure to consider the litter as the unit of analysis was also a key risk-of-bias concern. When there was not enough information to assess the potential bias for a risk-of-bias question and authors did not respond to an inquiry for further information, a conservative approach was followed, and the studies were rated probably high risk of bias for that question.

Table 4. OHAT Risk-of-bias Questions and Applicability by Study Design

Risk-of-bias Questions	Experimental Animal*	Human Controlled Trials**	Cohort	Case-Control	Cross-Sectional***	Case Series
1. Was administered dose or exposure level adequately randomized?	X	X				
2. Was allocation to study groups adequately concealed?	X	X				
3. Did selection of study participants result in the appropriate comparison groups?			X	X	X	
4. Did study design or analysis account for important confounding and modifying variables?			X	X	X	X
5. Were experimental conditions identical across study groups?	X					
6. Were research personnel blinded to the study group during the study?	X	X				
7. Were outcome data complete without attrition or exclusion from analysis?	X	X	X	X	X	
8. Can we be confident in the exposure characterization?	X	X	X	X	X	X
9. Can we be confident in the outcome assessment (including blinding of outcome assessors)?	X	X	X	X	X	X
10. Were all measured outcomes reported?	X	X	X	X	X	X
11. Were there no other potential threats to internal validity?	X	X	X	X	X	X

*Experimental animal studies are controlled exposure studies. Non-human animal observational studies can be evaluated using the design features of observational human studies such as cross-sectional study design.

**Human Controlled Trials are studies in humans with controlled exposure (e.g., Randomized Controlled Trials, non-randomized experimental studies)

***Cross-sectional studies include population surveys with individual data (e.g., NHANES) and surveys with aggregate data (i.e., ecological studies).

Any discrepancies in ratings between assessors were resolved by a senior technical specialist and through discussion when necessary to reach the final recorded risk-of-bias rating for each question along with a statement of the basis for that rating. Members of the evaluation team were consulted for assistance if additional expertise was necessary to reach final risk-of-bias ratings based on specific aspects of study design or performance reported for individual studies. Study procedures that were not reported were assumed not to have been conducted, resulting in an assessment of "probably high" risk of bias. Authors were queried by email to obtain missing information and responses received were used to update risk-of-bias ratings.

Table 5. The Four Risk-of-bias Rating Options	
Answers to the risk-of-bias questions result in one of the following four risk-of-bias ratings	
++	***Definitely Low*** *risk of bias*: There is direct evidence of low risk-of-bias practices
+	***Probably Low*** *risk of bias*: There is indirect evidence of low risk-of-bias practices **OR** it is deemed that deviations from low risk-of-bias practices for these criteria during the study would not appreciably bias results, including consideration of direction and magnitude of bias
– NR	***Probably High*** *risk of bias*: There is indirect evidence of high risk-of-bias practices (indicated with "-") **OR** there is insufficient information provided about relevant risk-of-bias practices (indicated with "NR" for not reported). Both symbols indicate probably high risk of bias.
––	***Definitely High*** *risk of bias*: There is direct evidence of high risk-of-bias practices

Organizing and Rating Confidence in Bodies of Evidence

Health Outcome Categories for Neurodevelopmental and Cognitive Effects

After data were extracted from all studies, the health effects results within the category of neurodevelopmental or cognitive effects were grouped across studies to develop bodies of evidence or collections of studies with data on the same or related outcomes. The grouping of health effect results was not planned a priori. The vast majority of the human studies evaluated intelligence quotient (IQ) in children as the single outcome; therefore, the discussion of cognitive neurodevelopmental effects in children focuses on IQ studies with supporting information from data on other endpoints. Cognitive function in adults was evaluated separately. Consistent with the NTP (2016) assessment, the primary focus within the animal study body of evidence was on animal studies with endpoints related to learning and memory.

Considerations for Pursuing a Narrative or Quantitative Evidence Synthesis

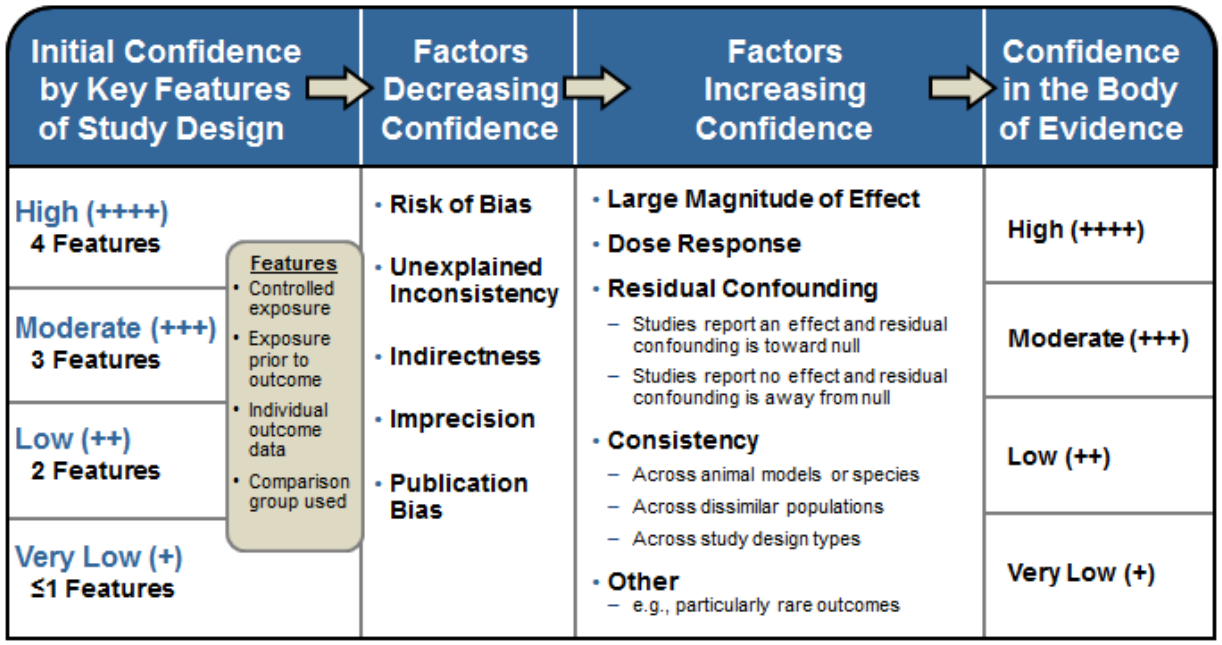
Heterogeneity within the available evidence was used to determine which type of evidence integration was appropriate—a quantitative synthesis (meta-analysis) or narrative approach for evidence integration. Choi *et al.* (2012) and Duan *et al.* (2018) conducted meta-analyses and found that high fluoride exposure was associated with lower IQ scores. Choi *et al.* (2012) was able to determine a risk ratio for living in an endemic fluorosis area but was unable to develop a dose-response relationship. Duan *et al.* (2018) suggested a significant non-linear dose-response relationship between fluoride dose and intelligence with the relationship stated to be most evident with exposures from drinking water

containing above 4 mg/L (or 4 ppm). Duan *et al.* (2018) found similar results as Choi *et al.* (2012) for the standardized mean difference; however, the majority of the available studies in both analyses compare populations with high fluoride exposure to those with lower fluoride exposure (with the lower exposure levels frequently in the range of drinking water fluoridation in the United States). After evaluating the available data, NTP determined that a narrative review—not a meta-analysis or other quantitative assessment—was appropriate for evidence integration due to heterogeneity in dose among the available human evidence, and because a hazard conclusion could be reached without conducting a meta-analysis. However, in the November 2019 review of the September 6, 2019 draft monograph (NASEM 2020), NASEM recommended that a meta-analysis be conducted. In response, NTP performed a meta-analysis of IQ studies in children. The meta-analysis protocol can be found with the revised systematic review protocol posted in September 2020 (https://ntp.niehs.nih.gov/go/785076).

Confidence Rating: Assessment of Body of Evidence

The quality of evidence for neurodevelopmental and cognitive function outcomes was evaluated using the GRADE system for rating the confidence in the body of evidence (Guyatt *et al.* 2011, Rooney *et al.* 2014). More detailed guidance on reaching confidence ratings in the body of evidence as "high," "moderate," "low," or "very low" is provided in the OHAT Handbook for Conducting a Literature-Based Health Assessment (http://ntp.niehs.nih.gov/go/38673, see STEP 5). In brief, available human and animal studies on a particular health outcome were initially grouped by key study design features, and each grouping of studies was given an initial confidence rating by those features. Starting at this initial rating (see column 1 of **Figure 1**), potential downgrading of the confidence rating was considered for factors that decrease confidence in the results (see column 2 of **Figure 1** [risk of bias, unexplained inconsistency, indirectness or lack of applicability, imprecision, and publication bias]); and potential upgrading of the confidence rating was considered for factors that increase confidence in the results (see column 3 of **Figure 1** [large magnitude of effect, dose response, consistency across study designs/populations/animal models or species, consideration of residual confounding, and other factors that increase our confidence in the association or effect]). Consideration of consistency across study designs, human populations, or animal species is not included in the GRADE guidance (Guyatt *et al.* 2011); however, it is considered in the modified version of GRADE used by OHAT (Rooney *et al.* 2014, NTP 2015).

Figure 1. Assessing Confidence in the Body of Evidence

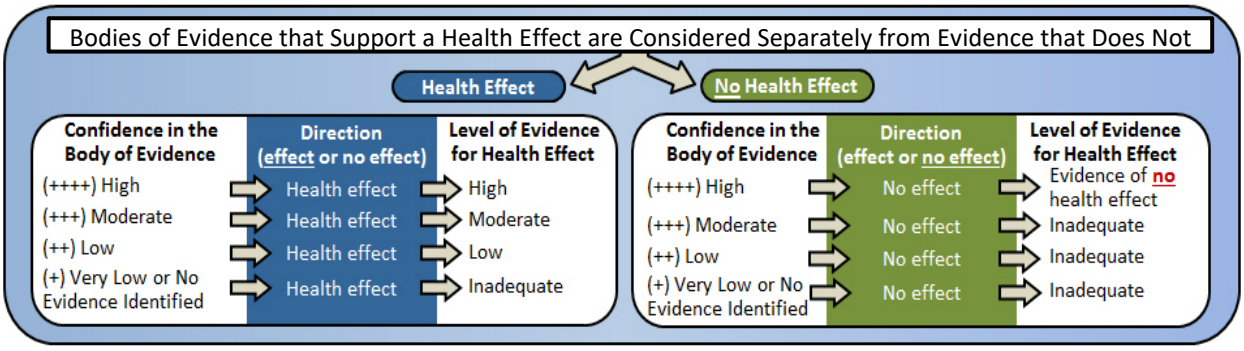


Confidence ratings were assessed by the evaluation team for accuracy and consistency, and discrepancies were resolved by consensus and consultation with technical advisors as needed. Confidence ratings for the primary outcomes are summarized in evidence profile tables for each outcome.

Preparation of Level of Evidence Conclusions

The confidence ratings were translated into level of evidence of health effects for each type of health outcome separately according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Inadequate (see **Figure 2**). The descriptor “evidence of no health effect” is used to indicate confidence that the substance is not associated with a health effect. Because of the inherent difficulty in proving a negative, the conclusion “evidence of no health effect" is only reached when there is high confidence in the body of evidence.

Figure 2. Translate Confidence Ratings into Evidence of Health Effect Conclusions



Evidence Descriptors	Definition
High Level of Evidence	There is high confidence in the body of evidence for an association between exposure to fluoride and the health outcome(s).
Moderate Level of Evidence	There is moderate confidence in the body of evidence for an association between exposure to fluoride and the health outcome(s).
Low Level of Evidence	There is low confidence in the body of evidence for an association between exposure to fluoride and the health outcome(s), or no data are available.
Inadequate Evidence	There is insufficient evidence available to assess if exposure to fluoride is associated with the health outcome(s).
Evidence of No Health Effect	There is high confidence in the body of evidence that exposure to fluoride is not associated with the health outcome(s).

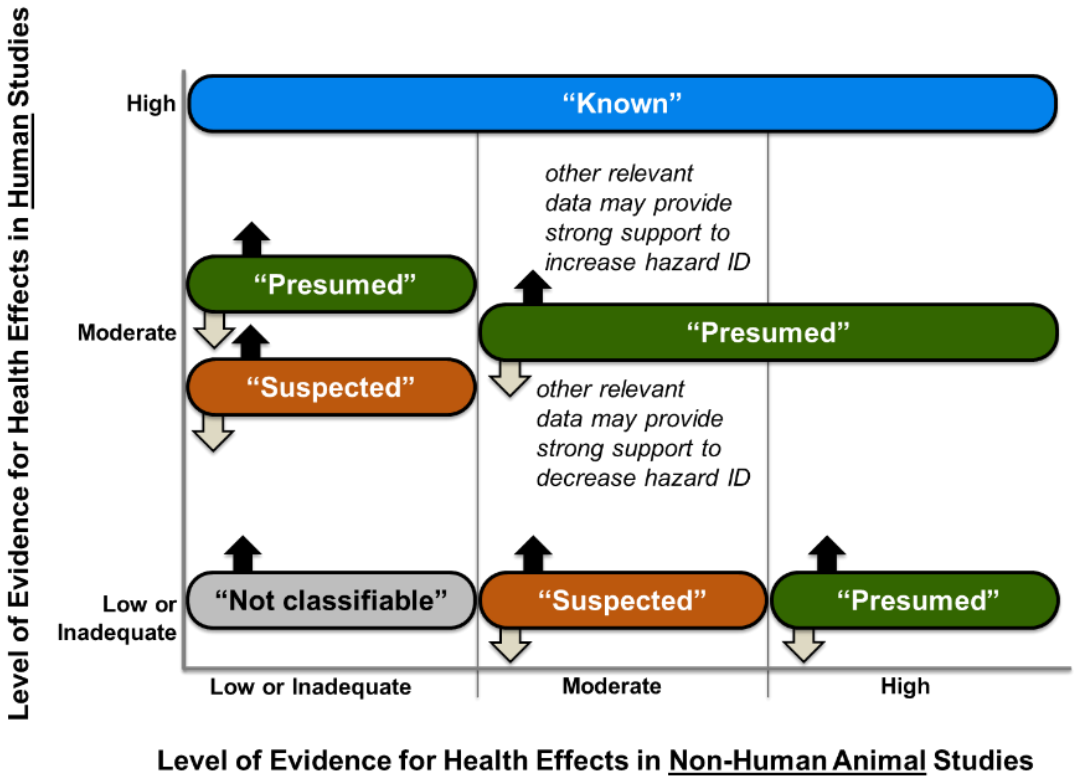
Integrate Evidence to Develop Hazard Identification Conclusions

Finally, the levels of evidence ratings for human and animal data were integrated with consideration of in vitro/mechanistic data to reach one of five possible hazard identification categories: (1) Known, (2) Presumed, (3) Suspected, (4) Not classifiable, or (5) Not identified to be a neurodevelopmental hazard to humans (see **Figure 3**).

Consideration of Human and Animal Data

Initial hazard identification conclusions were attempted by integrating the highest level-of-evidence conclusion for neurodevelopmental effects in children and cognitive effects in adults for the human and the animal evidence streams. The level-of-evidence conclusion for human data for neurodevelopmental or cognitive effects were considered with the level of evidence for non-human animal data to reach one of four initial hazard identification conclusions: Known, Presumed, Suspected, or Not classifiable. When either the human or animal evidence stream was characterized as "Inadequate Evidence," then conclusions were based on the remaining evidence stream alone (which is equivalent to treating the missing evidence stream as "Low" in **Figure 3**). If a moderate level-of-evidence conclusion for human data was reached with "Inadequate or Low Evidence" for the animal evidence stream, a hazard identification conclusion of either "suspected to be a hazard to humans" or "presumed to be a hazard to humans" could be reached based on scientific judgement as to the robustness of the body of evidence that supports moderate confidence in the human data and consideration of the potential impact of additional studies (NTP 2019).

Figure 3. Hazard Identification Scheme for Neurodevelopmental or Cognitive Effects



Consideration of Mechanistic Data

There is no requirement to consider mechanistic or mode-of-action data to reach a hazard identification conclusion regarding neurodevelopmental or cognitive health effects. However, when available, this and other relevant supporting types of evidence may be used to raise (or lower) the category of the hazard identification conclusion. Mechanistic data can come from a wide variety of studies that are not intended to identify a disease phenotype. This source of experimental data includes in vitro and in vivo laboratory studies directed at cellular, biochemical, genetic, and molecular mechanisms that attempt to explain how a chemical produces particular adverse effects.

For the evaluation of toxicity associated with fluoride exposure, NTP was interested in mechanistic or in vitro measures that comprise a coherent biological process that may support the plausibility of corresponding neurological outcomes reported from in vivo studies in animals or humans. The PECO statement in **Table 3** provides the specific endpoints considered including neuronal electrophysiology; mRNA, gene, or protein expression; cell proliferation or death in brain or thyroid tissue/cells; neuronal signaling; or synaptogenesis. In general, the mechanisms for fluoride-associated neurodevelopmental or cognitive effects are not well understood at this time, and mechanistic events identified in studies of animals receiving high fluoride exposures may not reflect biological processes occurring in humans at lower exposure levels. Mechanistic data from in vivo studies were used when feasible to examine the biological plausibility of the primary health outcomes considered in developing a hazard conclusion.

The factors outlined for increasing or decreasing confidence that the mechanistic data support biological plausibility are conceptually similar to those used to rate confidence in bodies of evidence for human or

animal in vivo studies are listed below and described in depth in the protocol (https://ntp.niehs.nih.gov/go/785076). Four factors were considered that contribute to increased confidence: potency, dose response, consistency in terms of cellular events observed at the same or lower doses than in vivo health effects, and consistency across cellular targets on the same functional pathway. Three factors were considered that contribute to decreased confidence: unexplained inconsistency across studies of the same endpoint, indirectness/applicability of the pathway for human health or concentrations for human exposure, and publication bias. Evaluations of the strength of evidence provided by mechanistic data were made on an outcome-specific basis based on discussion by the evaluation team and consultation with technical advisors as needed.

- If mechanistic data provided strong support for biological plausibility of the relationship between exposure and the health effect, the hazard identification conclusion may be upgraded (indicated by black "up" arrows in **Figure 3**) from that initially derived by considering the human and non-human animal evidence together.
- If mechanistic data provided strong opposition for biological plausibility of the relationship between exposure and the health effect, the hazard identification conclusion may be downgraded (indicated by gray "down" arrows in **Figure 3**) from that initially derived by considering the human and non-human animal evidence together.

Although it is envisioned that strong evidence for a relevant neurological effect from mechanistic data alone could indicate a potential that the substance is a neurodevelopmental hazard to humans, for this evaluation the mechanistic data were only considered to inform the biological plausibility of observed outcomes from in vivo exposure studies in humans or animals because of a general lack of understanding of the mechanistic basis for neurological outcomes.

RESULTS AND EVIDENCE SYNTHESIS

Literature Search Results

Literature Search Results Counts and Title and Abstract Screening

The electronic database searches retrieved 25,524 unique references in total (20,883 references during the initial search conducted in December 2016, 3,733 references during the literature search updates [including the final updated search conducted for the primary epidemiology studies on May 1, 2020], and 908 references from the supplemental Chinese database searches); 15 additional references were identified by technical advisors or from reviewing reference lists in published reviews and included studies. As a result of title and abstract screening, 1,038 references were moved to full-text review, 11,478 were excluded during manual title and abstract screening for not satisfying the PECO criteria, and an additional 13,023 were not screened and excluded based on the SWIFT algorithm.

Full-text Review

Among the 1,038 references that underwent full-text review, 499 references were excluded during the full-text review with reasons for exclusion documented at this stage; 332 references were excluded for not satisfying the PECO criteria; and 167 references from the May 2020 searches (main literature search update and supplemental Chinese database searches) were excluded for not including information that would materially advance the human, animal in vivo, or mechanistic findings (see the Literature Search Section for a description of the methodology). These screening results are outlined in a study selection diagram that reports numbers of studies excluded for each reason at the full text review stage (see **Figure 4**) [using reporting practices outlined in Moher *et al.* (2009)]. After full-text review, 539 studies were considered relevant with primary neurological outcomes, secondary neurological outcomes, and/or outcomes related to thyroid function (see **Appendix 2**). A few studies assessed data for more than one evidence stream (human, non-human mammal, and/or in vitro), and several human and animal studies assessed more than one type of outcome (e.g., primary and secondary outcomes). The number of included studies is summarized below. There are:

- 159 human studies (78 primary only; 13 secondary only; 5 primary and secondary; 6 primary and thyroid; 2 secondary and thyroid; and 55 thyroid only);
- 339 non-human mammal studies (7 primary only; 186 secondary only; 67 primary and secondary; 6 primary, secondary, and thyroid; 4 secondary and thyroid; and 69 thyroid only); and,
- 60 in vitro/mechanistic studies (48 neurological and 12 thyroid).

One publication contained human, experimental non-human mammal, and in vitro data. Three publications contained both human and experimental non-human mammal data. Fourteen publications contained data relevant to both experimental non-human mammal studies and in vitro studies.

Evaluation of SWIFT-Active Screener Results

During the initial title and abstract screening of 20,883 references using SWIFT-Active Screener, approximately 38%[6] of the 20,883 studies were manually screened in duplicate to identify an estimated 98.6% of the predicted number of relevant studies using the statistical algorithm in SWIFT-Active Screener (13,023 references were not screened). SWIFT-Active Screener predicted that there were 739 relevant studies during the initial title and abstract screening, of which 729 studies were identified and moved to full-text review. The SWIFT-Active statistical algorithm predicted that 10 relevant studies at the title and abstract level (10 represents 1.4% × 739 predicted relevant studies; or 739 predicted relevant studies minus 729 identified relevant studies during screening) were not identified by not screening the remaining 13,023 studies.

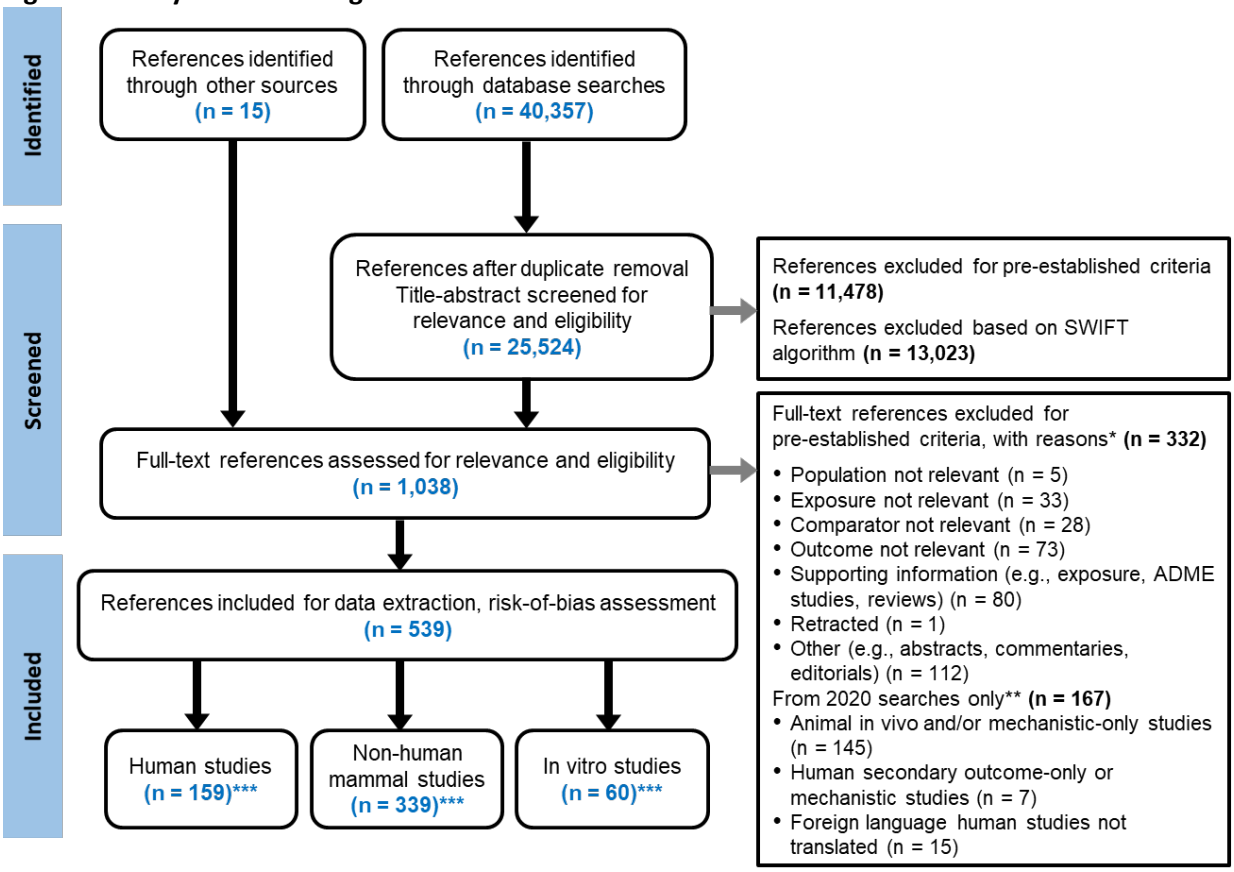
To further consider the impact of using SWIFT-Active Screener for this systematic review, NTP evaluated the SWIFT-Active screening results to gain a better understanding of the relevance of the last group of studies that were screened before 98% predicted recall was satisfied. The goal was to determine the likelihood of having missed important studies by not screening all of the literature. To do this, NTP evaluated subsets of studies screened in SWIFT-Active for trends and followed those studies through to full-text review for a final determination of relevance and potential impact (i.e., whether the studies had data on primary outcomes). Based on this evaluation, NTP estimates that the use of Swift-Active Screener may have resulted in missing 1–2 relevant human studies and 1–2 relevant animal studies with primary neurodevelopmental or cognitive outcomes.

Supplemental Chinese Database Searches and Human Epidemiology Studies

Following the NASEM committee peer review in November 2019 (NASEM 2020), supplemental searches were conducted in non-English language databases (CNKI and Wanfang). One focus of the screening of these supplemental search results was to identify null or no-effect studies that evaluated primary neurodevelopmental or cognitive outcomes (learning, memory, and intelligence) that may have been missed in previous approaches or may have been absent from the Fluoride Action Network website. Of the 908 references that were identified in the supplemental Chinese database searches, 16 relevant studies with primary neurological outcomes were identified (which were not identified through the main literature searches). Among these 16 studies, Kang *et al.* (2011) was the only null study with primary neurological outcomes that was identified through the supplemental Chinese database searches. NTP had the study translated to English, and the study was included. Note that Kang *et al.* (2011) is also identified by the Fluoride Action Network as a null study, but their website does not include an English translation of the study. The other 15 relevant studies contained results that would likely add to the body of evidence showing a negative association between fluoride exposure and primary neurological outcomes. Because this body of evidence is already so large, and because time was a factor in the revision of the monograph, these studies were not translated or included as this information would likely not materially advance the human findings.

[6] Howard *et al.* (2020) evaluated the performance of the SWIFT-Active Screener methods for estimating total number of relevant studies using 26 diverse systematic review datasets that were previously screened manually by reviewers. The authors found that on average, 95% of the relevant articles were identified after screening 40% of the total reference list when using SWIFT-Active Screener. In the document sets with 5,000 or more references, 95% of the relevant articles were identified after screening 34% of the available references, on average, using SWIFT-Active Screener.

Figure 4. Study Selection Diagram



* Studies may have been excluded for more than one reason; the first reason identified by the screener was recorded.

** Animal in vivo, human secondary outcome-only, and human and animal mechanistic references from the 2020 database searches were scanned for evidence that might strengthen the information in the September 6, 2019 draft monograph. Although 145 additional animal in vivo and/or mechanistic studies and 7 additional human secondary outcome-only or mechanistic-only studies were identified, information that would materially advance the human, animal in vivo, and mechanistic findings was not identified; therefore, these studies were not included. Additionally, 15 human primary outcome studies from the 2020 Chinese database search were excluded based on English abstracts and google translations because information that would materially advance the human findings was not identified; 1 null publication from the 2020 Chinese database search (Kang *et al.* 2011) was identified, translated, and included.

*** One publication contained human, experimental non-human mammal, and in vitro data. Three publications contained both human and experimental non-human mammal data. Fourteen publications contained data relevant to both experimental non-human mammal studies and in vitro studies.

Neurodevelopmental and Cognitive Health Effects Results

All the neurodevelopmental and cognitive data were initially considered and evaluated, with more in-depth analysis where similar endpoints were evaluated across multiple studies (e.g., IQ). Hazard conclusions were developed separately for two different age groups (i.e., children and adults) to address potential differences in the health impact based on exposure during development compared to adulthood. Although the data cover a wide array of endpoints (see **Figure 5**), the hazard conclusion covers a single category for each age group. The largest bodies of evidence were for IQ (n = 71 studies), learning and memory (n = 8 studies), as well as other cognitive development effects (e.g., total

neurobehavioral scores and total mental capacity index in children and cognitive impairment in adults; n = 14 studies)[7]. Due to heterogeneity in the endpoints examined and the limited number of human or animal studies, congenital neurological malformations and neurological complications of fluorosis were not evaluated because the body of evidence was inadequate to evaluate these potential effects. These health outcomes are not further discussed in this assessment. To the extent possible, human and animal data were grouped into similar categories (e.g., IQ in humans was considered comparable to learning and memory in animals). NTP had previously assessed animal data related to effects on learning and memory associated with fluoride exposure (NTP 2016). Therefore, to update the conclusions of the NTP (2016) systematic review, only more recent animal studies were evaluated in this assessment. Although the previous NTP (2016) report was conducted through January 14, 2016, the current assessment included studies published from 2015 through April 2019 and considered studies from the NTP (2016) report. Thirty-five animal studies have been identified that met these criteria, including 23 studies with learning and memory endpoints and 12 studies with only motor and sensory endpoints. Consistent with the NTP (2016) assessment, only learning and memory studies have been considered in the development of hazard identification conclusions. The additional motor and sensory studies have been considered, along with information on motor and sensory effects reported in the learning and memory studies, to provide evidence of possible indirectness related to the learning and memory assessments.

Risk-of-bias Considerations

Risk-of-bias ratings for each individual study for all risk-of-bias questions are available in **Appendix 3**. The risk of bias of individual studies in the body of evidence was considered in developing confidence ratings. The key risk-of-bias questions (i.e., confounding, exposure characterization, and outcome assessment for human studies and randomization, exposure characterization, and outcome assessment for experimental animal studies) are discussed in the consideration of the body of evidence. For this assessment, the key risk-of-bias questions, if not addressed appropriately, are considered to potentially have the greatest impact on the results. In addition, for developmental studies in animals, controlling for potential litter effects (i.e., adjusting for similarities in responses between littermates) was also a key risk-of-bias concern. The other risk-of-bias questions were also taken into consideration and were used to identify any other risk-of-bias concerns that may indicate serious issues with the studies. No study was excluded based on concerns for risk of bias; however, confidence conclusions were considered with and without higher risk-of-bias studies (i.e., studies rated probably high risk of bias for at least two key risk-of-bias questions or definitely high for any single question) to assess the impact of the higher risk-of-bias studies. The remaining studies (i.e., other than the higher risk-of-bias studies) were considered lower risk of bias. Based on NASEM recommendations (NASEM 2020), **Appendix 4** was created for the lower risk-of-bias studies to describe strengths and limitations of the studies identified during the assessment and to clarify why they are considered to pose lower risk of bias.

Human Neurodevelopmental and Cognitive Data

While there were several neurodevelopmental and cognitive endpoints assessed (see **Figure 5**), most of the available studies evaluated intelligence (e.g., IQ) in children. Other measures of neurodevelopment or cognitive function in children were also assessed, including general cognitive index (GCI), mental capacity, mental development index (MDI), or neonatal behavioral neurological assessment (NBNA). However, because the majority of studies evaluated intelligence, the discussion focuses primarily on IQ in children with separate discussions on other measures of cognitive function and neurobehavioral

[7]Some studies are included in more than endpoint category (e.g., IQ and other cognitive developmental effects); therefore, these counts are not mutually exclusive.

effects in children and cognitive effects in adults. The available body of literature that evaluates the association between fluoride exposure and neurodevelopmental and cognitive effects is relatively robust (n = 92) and confidence considerations in the body of evidence and hazard conclusions are focused on the studies with the least potential for bias (n = 31). Studies with higher potential for bias (n = 61) have also been evaluated and determined to have little impact on the confidence and hazard conclusions. A subgroup analysis within the meta-analysis (described in detail in **Appendix 5**) demonstrated that results were robust to the exclusion of higher risk-of-bias studies (see **Appendix 5**, **Figure A5-6**). All evaluated studies can be found in **Appendix 2**.

This section is organized to present and explain NTP's two confidence ratings in the bodies of evidence from epidemiological studies that fluoride exposure is associated with cognitive neurodevelopmental effects in children and cognitive effects in adults. These confidence ratings were determined as described in **Figure 1**.

Summary: There is moderate confidence in the body of evidence that fluoride exposure is associated with cognitive neurodevelopmental effects in children, and low confidence in the body of evidence that fluoride exposure is associated with cognitive effects in adults. The moderate confidence rating is supported by consistent evidence from the available studies of an association between high-fluoride exposure (mainly greater than the WHO Drinking Water Quality Guideline [>1.5 mg/L] (WHO 2011), but also high exposure via fluoridated salt and food) and lower IQ or cognitive function in children. There is also a recent study of lower IQ in children living in areas where drinking water fluoride concentrations are <1.5 mg/L. Specifically, a study conducted in Canada observed significantly lower IQ scores in boys and girls associated with higher estimated total maternal consumption of fluoride during pregnancy from drinking water and other water-based beverages including black and green tea. When looking at maternal urinary fluoride concentrations, the significant negative association with IQ scores was seen in in boys but not girls (Green *et al.* 2019). Another study conducted in Mexico with similar maternal urinary fluoride concentrations during pregnancy as seen in Green *et al.* (2019) observed significantly lower IQ scores in boys and girls associated with higher perinatal exposure to fluoride (Bashash *et al.* 2017). Although the body of evidence in children supports lower IQ with increased fluoride exposure, there is a lack of evidence of an association between exposure to fluoride and cognitive effects in adults (Jacqmin *et al.* 1994, Li *et al.* 2016). The body of evidence available to examine the association between exposure to fluoride and cognitive effects in adults is limited to two lower risk-of-bias cross-sectional studies; due to the limited number of studies and a lack of an observed effect, this body of evidence is considered inadequate to evaluate whether fluoride exposure is associated with cognitive effects in adults (see **Table 7**).

Most of the available epidemiological studies that evaluated the association between fluoride exposure and cognitive neurodevelopmental effects assessed IQ and other measures of cognitive function in children (see **Figure 5**). Confidence conclusions are based on those studies with the lowest potential for bias (n = 28; 26 in children and 2 in adults) (see **Table 6**). Most of these studies measured fluoride levels in drinking water or urine. All but two of the studies were conducted in infants or children. The two studies in adults were conducted in older adult populations (≥60 years old; one in France and the other in a fluorosis-endemic area of China) to evaluate the effects of fluoride on cognitive impairment.

The studies in children were conducted in multiple populations. Of the 26 studies in children:

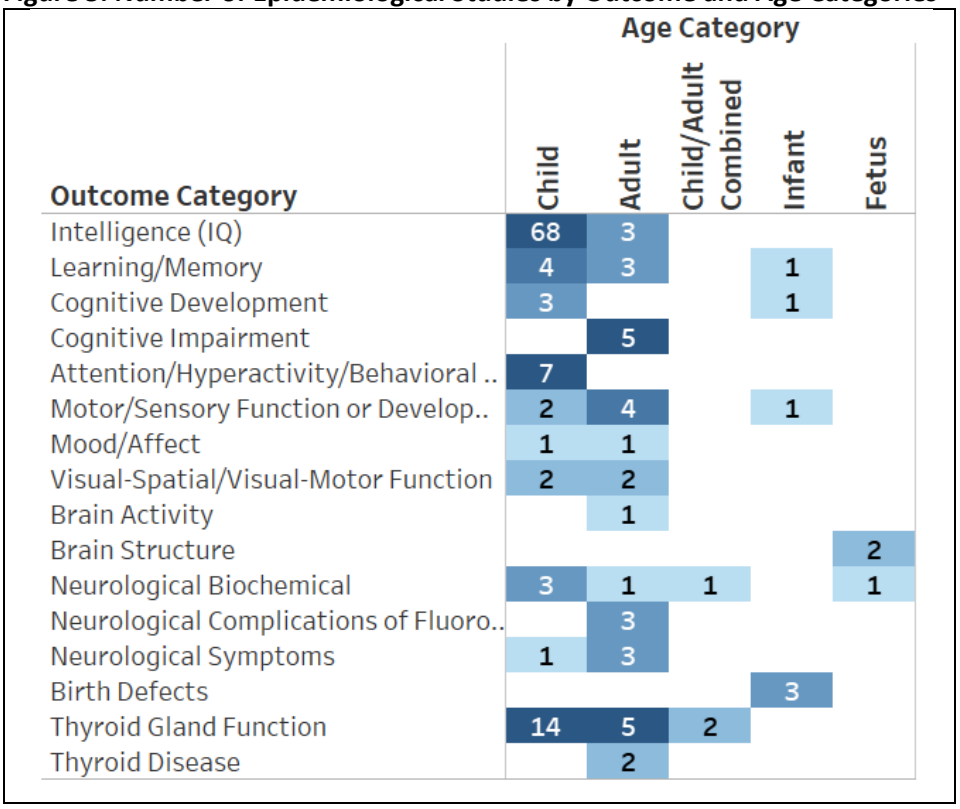
- 12 were conducted in 6 areas of China based on 8 study populations (1 study with both IQ and other neurodevelopmental outcomes, 9 studies with IQ only, and 2 studies with other neurodevelopmental outcomes);
- 6 were conducted in 4 areas of Mexico based on 5 study populations (1 study with both IQ and other neurodevelopmental outcomes, 2 studies with IQ only, and 3 studies with other neurodevelopmental outcomes);
- 4 were conducted in Canada using 2 separate cohorts (2 studies with IQ only and 2 studies with other neurodevelopmental outcomes);
- 3 were conducted in 3 areas of India (all IQ studies); and
- 1 was conducted in Iran (IQ study).

The IQ studies used many different tests to measure IQ. The IQ tests used often differed by population as not all IQ tests are appropriate for all populations (e.g., western vs. Asian populations). In some cases, different IQ tests were used to study similar populations. Overall, studies used IQ or cognitive tests appropriate for the population and were age appropriate. Other neurodevelopmental outcomes assessed in some studies included neurobehavioral effects (in infants), learning and memory impairment, and learning disabilities such as attention deficit hyperactivity disorder (ADHD). The different tests conducted and the populations on which the tests were conducted are indicated in **Table 6**.

The lower risk-of-bias studies (i.e., studies not meeting criteria for higher risk of bias) showing associations with cognitive neurodevelopmental effects in children include 5 prospective cohort studies from 3 study populations (Bashash *et al.* 2017, Valdez Jimenez *et al.* 2017, Green *et al.* 2019, Bashash *et al.* 2018, Till *et al.* 2020) and 21 cross-sectional studies from 16 study populations (Li *et al.* 2004 [translated in Li *et al.* 2008a], Choi *et al.* 2015, Rocha-Amador *et al.* 2007, Rocha-Amador *et al.* 2009, Saxena *et al.* 2012, Seraj *et al.* 2012, Xiang *et al.* 2003a, Xiang *et al.* 2011, Zhang *et al.* 2015b, Ding *et al.* 2011, Barberio *et al.* 2017b, Yu *et al.* 2018, Cui *et al.* 2018, Cui *et al.* 2020, Wang *et al.* 2020b, Wang *et al.* 2020a, Wang *et al.* 2012, Soto-Barreras *et al.* 2019, Sudhir *et al.* 2009, Trivedi *et al.* 2012, Riddell *et al.* 2019) (see **Figure D1** through **Figure D12**). One limitation of the 21 cross-sectional studies was the lack of direct evidence that exposure to fluoride occurred prior to the development of the neurodevelopmental outcomes. However, several studies from different study populations (n = 5) indicated that a large portion of the exposed children had dental fluorosis (ranging from 43–100%) at the time of the assessment (Choi *et al.* 2015, Ding *et al.* 2011, Seraj *et al.* 2012, Yu *et al.* 2018, Sudhir *et al.* 2009). Because dental fluorosis occurs when fluoride is consumed during enamel formation usually during the first 6–8 years of life, the presence of dental fluorosis suggests that exposures to fluoride occurred prior to the outcome assessment. Ten studies from seven study populations (including Yu *et al.* (2018), Wang *et al.* (2012) listed above) excluded subjects that had not lived in the study area for a specified period of time, sometimes since birth (Rocha-Amador *et al.* 2007, Rocha-Amador *et al.* 2009, Saxena *et al.* 2012, Yu *et al.* 2018, Wang *et al.* 2020b, Wang *et al.* 2012, Xiang *et al.* 2011, Xiang *et al.* 2003a, Soto-Barreras *et al.* 2019, Sudhir *et al.* 2009). Another study evaluated fluoride exposure in mothers and included urine levels just prior to birth and assessed children a few days after birth (Li *et al.* 2004 [translated in Li *et al.* 2008a]). Because these areas were generally known to be fluoride-endemic areas for long periods of time, it can generally be assumed that in these 14 cross-sectional studies from 11 study populations, exposure occurred prior to the outcome. These exposure concerns were not an issue for the prospective studies because fluoride levels were measured prenatally. Therefore, the moderate confidence in the body of evidence in children is primarily based on the consistency of findings across different populations in the 5 lower risk-of-bias prospective cohort studies and the 14

cross-sectional studies where exposure is considered to have occurred prior to the outcome with initial and final ratings of moderate confidence.

Figure 5. Number of Epidemiological Studies by Outcome and Age Categories*



*Interactive figure and additional study details in (https://public.tableau.com/profile/ntp.visuals#!/vizhome/Fluoride_Epi_2020Update/Figure5)

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Children-IQ Studies					
China					
Choi *et al.* (2015)	Cross-sectional Mianning County/1st grade children [51]	Drinking water GM: 2.20 mg/L Children's urine GM: 1.64 mg/L Severity of fluorosis (Dean Index)	Children (ages 6–8 years)	IQ: WISC-IV (square root block design and digit span)	Compared to normal/ questionable fluorosis, moderate/severe fluorosis significantly associated with lower total (adjusted β = −4.28; 95% CI: −8.22, −0.33) and backward digit span scores (adjusted β = −2.13; 95% CI: −4.24, −0.02); linear correlation between fluoride in urine (adjusted β = −1.67; 95% CI: −5.46, 2.12) and in drinking water (adjusted β = −1.39; 95% CI: −6.76, 3.98) with total digit span was observed but not significant; other outcomes not significantly associated with fluoride exposure Adjusted for child's age, child's gender, parity, illness before 3 years old, household income last year, and caretaker's age and education
Cui *et al.* (2018)	Cross-sectional Tianjin City (districts Jinghai and Dagang)/school children [323]	Children's urine Range (log-transformed): −1.2–2.2	Children (ages 7–12 years)	IQ: Combined Raven's Test for Rural China	Significant correlation between IQ score and urinary fluoride (adjusted β = −2.47) Adjusted for child age, mother's education, family member smoking, stress, and anger
Cui *et al.* (2020)	Cross-sectional Tianjin City (all districts) /school children [498]	Children's urine <1.6–≥2.5 mg/L	Children (ages 7–12 years)	IQ: Combined Raven's Test	No significant difference in IQ score in the three urinary fluoride exposure groups based on a one-way ANOVA <1.6 mg/L = 112.16 ± 11.50 1.6-2.5 mg/L = 112.05 ± 12.01 ≥2.5 mg/L = 110 ± 14.92 No statistical adjustment for confounders

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Zhang *et al.* (2015b)	Cross-sectional Tianjin City (Jinnan District)/school children [180]	Drinking water Mean: 0.63 (control), 1.40 (endemic fluorosis) mg/L (SD not reported) Children's urine Mean (SD): 1.1 (0.67) (control), 2.4 (1.01) (endemic fluorosis) mg/L Children's serum Mean (SD): 0.06 (0.03) (control), 0.18 (0.11) (endemic fluorosis) mg/L	Children (ages 10–12 years)	IQ: Combined Raven's Test for Rural China	Significant correlation between IQ score and serum fluoride (r = −0.47) and urinary fluoride (r = −0.45); significant difference in IQ score for high-fluoride area (>1 ppm; 102.33 ± 13.46) compared with control area (109.42 ± 13.30) Adjusted for child's age and gender, if applicable
Yu *et al.* (2018)*	Cross-sectional Tianjin City (7 towns) /children [2,886]	Drinking water Mean (SD): 0.50 (0.27) (normal), 2.00 (0.75) (high) mg/L Children's urine Mean (SD): 0.41 (0.49) (normal), 1.37 (1.08) (high) mg/L	Children (ages 7–13 years)	IQ: Combined Raven's Test for Rural China	Significant difference in mean IQ scores in high water fluoride area (>1.0 mg/L; 106.4 ± 12.3 IQ) compared to the normal area (≤1.0 ppm; 107.4 ± 13.0 IQ); distribution of the IQ scores also significantly different (p = 0.003); every 0.5-mg/L increase in water fluoride was associated with a 4.29 lower IQ score (95% CI: −8.09, −0.48) between 3.40 and 3.90 mg/L Adjusted for child's age, child's gender, maternal education, paternal education, and low birth weight
Wang *et al.* (2020b)	Cross-sectional Tianjin City (villages not specified)/school children [571]	Drinking water Mean (SD): 1.39 (1.01) mg/L Children's urine Mean (SD): 1.28 (1.30) mg/L	Children (ages 7–13 years)	IQ: Combined Raven's Test for Rural China	Significant associations between IQ and water and urine fluoride concentrations in boys and girls combined based on both quartiles and continuous measures (water: −1.587 per 1-mg/L increase; urine: −1.214 per 1-mg/L increase); there was no significant modification effect of gender Adjusted for child's age, child's gender, BMI, maternal education, paternal education, household income, and low birth weight

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Ding *et al.* (2011)	Cross-sectional Inner Mongolia (Hulunbuir City)/ elementary school children [331]	Drinking water Mean (SD): 1.31 (1.05) mg/L Children's urine Range: 0.1–3.55 mg/L	Children (ages 7–14 years)	IQ: Combined Raven's Test for Rural China	Significant association between urinary fluoride and IQ score (each 1-mg/L increase was associated with a lower IQ score of 0.59 points (95% CI: −1.09, −0.08); dose response relationship between fluoride and dental fluorosis ($p < 0.0001$) Adjusted for child's age
Xiang *et al.* (2003a)	Cross-sectional Wamiao and Xinhuai villages (Sihong County)/school children [512]	Drinking water Mean (SD): 0.36 (0.15) (control), 2.47 (0.79) (high fluoride) mg/L Children's urine Mean (SD): 1.11 (0.39) (control), 3.47 (1.95) (high fluoride) mg/L Village of residence (non-endemic v. endemic fluorosis)	Children (ages 8–13 years)	IQ: Combined Raven's Test for Rural China	Significant dose-related effect of fluoride on IQ score based on quintile levels with significantly lower IQ scores observed at water fluoride levels 1.53 mg/L or higher; Pearson correlation coefficient of -0.164 with urinary fluoride; IQ scores for children in non-endemic region (100.41 ± 13.21) significantly higher than endemic region (92.02 ± 13.00); calculated a lower-bound confidence limit benchmark concentration (BMCL) of 1.85 mg/L
Xiang *et al.* (2011)	Cross-sectional Wamiao and Xinhuai villages (Sihong County)/school children (same population as Xiang *et al.* (2003a)) [512]	Children's serum Mean (SD): 0.041 (0.009) (control), 0.081 (0.019) (high fluoride) mg/L	Children (ages 8–13 years)	IQ: Combined Raven's Test for Rural China	Significant trend on association between quartiles of serum fluoride and children's IQ score < 80 (adjusted ORs for Q1 and Q2; Q1 and Q3; and Q1 and Q4, respectively: 1; 2.22 [95% CI: 1.42, 3.47]; and 2.48 [95% CI: 1.85, 3.32]); significant effects at ≥0.05 ppm fluoride Adjusted for child's age and gender
Wang *et al.* (2012)	Cross-sectional Wamiao and Xinhuai villages (Sihong County)/school children (same population as Xiang *et al.* (2003a)) [526]	Drinking water Mean (SD): 0.36 (0.11) (control), 2.45 (0.80) (high fluoride) mg/L Children's total fluoride intake Mean (SD): 0.78 (0.13) (control), 3.05 (0.99) (high fluoride) mg/day Village of residence (non-endemic v. endemic fluorosis)	Children (ages 8–13 years)	IQ: Combined Raven's Test for Rural China	Significantly lower mean IQ in the high fluoride village (92.02 ± 13.00) compared to the control village (100.41 ± 13.21); when high exposure group was broken into 4 exposure groups, a dose-dependent decreasing IQ and increase in % with low IQ observed; significant correlation between total fluoride intake and IQ ($r = −0.332$); OR for IQ<80 per increase in total fluoride intake = 1.106; 95% CI 1.052–1.163). Adjusted for child's age and gender

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Mexico					
Rocha-Amador *et al.* (2007)	Cross-sectional Moctezuma and Salitral in San Luis Potosi State and 5 de Febrero of Durango State /elementary school children [132]	Drinking water Mean (SD): 0.8 (1.4), 5.3 (0.9), 9.4 (0.9) mg/L (3 rural areas) Children's urine Mean (SD): 1.8 (1.5), 6.0 (1.6), 5.5 (3.3) mg/L (3 rural areas)	Children (ages 6–10 years)	IQ: WISC-Revised Mexican Version	Significant associations between fluoride and IQ scores (full IQ adjusted βs of −10.2 with water and −16.9 with urine; CIs not reported); arsenic also present, but the effect was smaller (full IQ adjusted βs of −6.15 with water and −5.72 with urine; CIs not reported) Adjusted for blood lead, mother's education, SES, height-for-age z-scores, and transferrin saturation
Bashash *et al.* (2017)	Cohort (prospective) Mexico City/Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) participants [299] IQ analysis [211]	Maternal urine during pregnancy Mean (SD): 0.90 (0.35) mg/L Children's urine Mean (SD): 0.82 (0.38) mg/L	Children (ages 6–12 years)	IQ: WASI-Spanish Version	Significant effect between maternal urinary fluoride and offspring IQ score (adjusted β = −2.50; 95% CI: −4.12, −0.59); associations with children's urine not significant Adjusted for gestational age, weight at birth, child's gender, parity (being the first child), age at outcome measurement, and maternal characteristics including smoking history (ever smoked during the pregnancy vs. nonsmoker), marital status (married vs not married), age at delivery, education, IQ, and cohort

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Soto-Barreras *et al.* (2019)	Cross-sectional Chihuahua/school children [161]	Drinking water Range: 0.05–2.93 mg/L Children's urine Range: 0.11–2.10 mg/L	Children (ages 9–10 years)	IQ: Raven's Colored Progressive Matrices	No significant differences in fluoride exposure level (urine fluoride [p = 0.559], exposure dose [p = 0.389], or fluorosis index [p = 0.851]) between the different IQ grades No statistical adjustment for confounders
Canada					
Green *et al.* (2019)	Cohort (prospective) 10 cities/Maternal-Infant Research on Environmental Chemicals (MIREC) [512] Non-Fluoridated [238] Fluoridated [162] Boys [248] Girls [264]	Maternal urine during pregnancy Mean (SD): 0.51 (0.36) mg/L (0.40 [0.27] mg/L in non-fluoridated areas and 0.69 [0.42] mg/L in fluoridated areas) Maternal fluoride intake during pregnancy Mean (SD): 0.54 (0.44) mg/day (0.30 [0.26] and 0.93 [0.43] mg/day, respectively) Drinking water Mean (SD): 0.31 (0.23) mg/L (0.13 [0.06] and 0.59 [0.08] mg/L, respectively)	Children (ages 3−4 years)	IQ: full scale, performance, and verbal using Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III)	Significantly lower full-scale IQ (adjusted β = −4.49; 95% CI: −8.38, −0.60) and performance IQ (adjusted β = −4.63; 95% CI: −9.01, −0.25) per 1-mg/L increase in maternal urine in boys, but not girls (adjusted β = 2.40; 95% CI: −2.53, 7.33 and adjusted β = 4.51; 95% CI: −1.02, 10.05, respectively); significantly lower full-scale IQ (adjusted β = −3.66; 95% CI: −7.16, −0.15) per 1-mg increase in maternal fluoride intake (no sex interaction); significantly lower full-scale IQ (adjusted β = −5.29; 95% CI: −10.39, −0.19) per 1-mg/L increase in water fluoride concentration (no sex interaction); no significant changes observed in verbal IQ Adjusted for city, HOME score, maternal education, race, child's gender, and prenatal secondhand smoke exposure

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Till *et al.* (2020)	Cohort (prospective) 10 cities/ MIREC [398] Non-Fluoridated [247] Fluoridated [151] Breastfed as infants [200] Formula-fed as infants [198]	Maternal urine during pregnancy Mean (SD) breastfed: 0.42 (0.28) mg/L in non-fluoridated areas and 0.70 (0.39) mg/L in fluoridated areas formula-fed: 0.38 (0.27) mg/L in non-fluoridated areas and 0.64 (0.37) mg/L in fluoridated areas Infant fluoride intake Mean (SD) breastfed: 0.02 (0.02) mg/day in non-fluoridated areas and 0.12 (0.07) mg/day in fluoridated areas formula fed: 0.08 (0.04) mg/day in non-fluoridated areas and 0.34 (0.12) mg/day in fluoridated areas Drinking water Mean (SD) breastfed: 0.13 (0.06) mg/L in non-fluoridated areas and 0.58 (0.08) mg/L in fluoridated areas formula fed: 0.13 (0.05) mg/day in non-fluoridated areas and 0.59 (0.07) mg/L in fluoridated areas	Children (ages 3–4 years)	IQ: full-scale, performance, and verbal using Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III)	Significantly lower performance IQ with water fluoride (−9.26 formula-fed, −6.19 breastfed) and fluoride intake from formula (−8.76); significantly lower full scale IQ with water fluoride in formula-fed (−4.40); lower full-scale IQ for water fluoride in breastfed (−1.34) and fluoride intake from formula (−2.69) were not significant; no significant changes in verbal IQ scores with fluoride exposure Adjusted for maternal education, maternal race, child's age at IQ testing, child's sex, HOME total score, and second-hand smoke status in the child's house (separate analysis also adjusted for mother's urinary fluoride)
India					
Sudhir *et al.* (2009)	Cross-sectional Nalgonda District (Andhra Pradesh)/school children [1,000]	Drinking water Level 1: <0.7 ppm Level 2: 0.7–1.2 ppm Level 3: 1.3–4.0 ppm Level 4: >4.0 ppm	Children (ages 13–15 years)	IQ: Raven's Standard Progressive Matrices	Significant increase in mean and distributions of IQ grades (i.e., increase in intellectually impaired children) with increasing drinking water fluoride levels No statistical adjustment for confounders

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Saxena *et al.* (2012)	Cross-sectional Madhya Pradesh/school children [170]	Drinking water ≥1.5 mg/L (high fluoride group) Children's urine Range: 1.7–8.4 mg/L	Children (age 12 years)	IQ: Raven's Standard Progressive Matrices	Significant correlation between water (r = 0.534; p = 0.000) and urinary fluoride (r = 0.542; p = 0.000) levels and IQ score; no significant differences in the levels of urinary lead or arsenic in children from the different groups Confounders included in the analysis were not reported
Trivedi *et al.* (2012)	Cross-sectional Kachchh, Gujarat/school children (6th and 7th grades) [84]	Drinking water Mean (SE): 0.84 (0.38) (low), 2.3 (0.87) (high) Children's urine Mean (SE): 0.42 (0.23) (low), 2.69 (0.92) (high)	Children (age 12–13 years)	IQ: questionnaire prepared by Professor JH Shah (97% reliability rating)	Significantly lower IQ score in the high fluoride (92.53 ± 3.13) compared to the low fluoride (97.17 ± 2.54) areas in boys and girls combined (as well as separately) No statistical adjustment for confounders
Iran					
Seraj *et al.* (2012)	Cross-sectional Makoo/school children [293]	Drinking water Mean (SD): 0.8 (0.3) (normal), 3.1 (0.9) (medium), 5.2 (1.1) (high) mg/L	Children (ages 6–11 years)	IQ: Raven's Colored Progressive Matrices	Significant correlation between water fluoride and IQ score (adjusted β = −3.865; CIs not reported); significantly higher IQ score in normal area (97.77 ± 18.91) compared with medium (89.03 ± 12.99) and high (88.58 ± 16.01) areas Adjusted for child's age, child's gender, child's education level, mother's education level, father's education level, and fluorosis intensity
Children-Other Neurodevelopmental Studies					
China					
Choi *et al.* (2015)	Cross-sectional Mianning County/1st grade children [51]	Drinking water GM: 2.20 mg/L Children's urine GM: 1.64 mg/L Severity of fluorosis (Dean Index)	Children (ages 6–8 years)	Learning and memory: Neuropsychological tests including WRAML Visual motor ability: WRAVMA Motor ability: Finger tapping task Manual dexterity: Grooved pegboard test	Outcomes unrelated to the IQ test not significantly associated with fluoride exposure Adjusted for child's age, child's gender, parity, illness before 3 years old, household income last year, and caretaker's age and education

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Li *et al.* (2004) [translated in Li *et al.* 2008a]	Cross-sectional Zhaozhou County, Heilongjiang Province/neonates [91]	Drinking water Range: 0.5–1.0 mg/L (control); 1.7–6.0 mg/L (high) Maternal urine during pregnancy Mean (SD): 1.74 (0.96) mg/L (control); 3.58 (1.47) mg/L (high)	Neonates (24–72 hours after delivery)	Neurodevelopmental: Neonatal behavioral neurological assessment (NBNA)	Significant differences in neurobehavioral assessment total scores between high-fluoride (36.48 ± 1.09) and control groups (38.28 ± 1.10); significant differences in total score of behavioral capability that includes measures of non-biological visual orientation reaction and biological visual and auditory orientation reaction between the two groups (11.34 ± 0.56 in controls compared to 10.05 ± 0.94 in high-fluoride group) No statistical adjustment for confounders
Wang *et al.* (2020a)	Cross-sectional Tongxu County/school children [325]	Children's urine Mean (SD): 1.54 (0.89) mg/L	Children (ages 7–13 years)	ADHD and behavior measures: Conners' Parent Rating Scale-Revised (Chinese version) (CPRS-48)	Significant association between psychosomatic problems and urinary fluoride level (per 1-mg/L increase β=4.01; 95% CI 2.74, 5.28; OR for T-score >70=1.97; 95% CI 1.19, 3.27) Adjusted for child's age, child's gender, child's BMI, urinary creatinine, mother migrated and father migrated
Mexico					
Rocha-Amador *et al.* (2009)	Cross-sectional Durango/elementary school children [80]	Children's urine GM (SD): 5.6 (1.7) mg/L	Children (ages 6–11 years)	Visuospatial organization and visual memory: Rey-Osterrieth Complex Figure Test, children's version	Significant correlation between urinary fluoride and visuospatial organization (r = −0.29) and visual memory scores (r = −0.27); no significant correlation with arsenic Adjusted for age
Valdez Jimenez *et al.* (2017)	Cohort (Prospective) Durango City and Lagos de Moreno/infants [65]	Drinking water Range: 0.5–12.5 mg/L (all trimesters) Maternal urine Range: 0.16–8.2 mg/L (all trimesters)	Infants (ages 3–15 months)	Mental development index (MDI): Bayley Scales of Infant Development II (BSDI-II)	Significant correlation between maternal urinary fluoride and MDI score during first trimester (adjusted β = −19.05; SE = 8.9) and second trimester (adjusted β = −19.34; SE = 7.46) Adjusted for gestational age, child's age, marginality index, and type of drinking water

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Bashash *et al.* (2017)	Cohort (prospective) Mexico City/Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) participants [299] GCI analysis [287]	Maternal urine during pregnancy Mean (SD): 0.90 (0.35) mg/L Children's urine Mean (SD): 0.82 (0.38) mg/L	Children (age 4 years)	General cognitive index (GCI): McCarthy Scales of Children's Abilities (MSCA)	Significant effect between maternal urinary fluoride and offspring GCI score (adjusted β = −3.15; 95% CI: −5.42, −0.87); associations with children's urine not significant Adjusted for gestational age, weight at birth, child's gender, parity (being the first child), age at outcome measurement, and maternal characteristics including smoking history (ever smoked during the pregnancy vs. nonsmoker), marital status (married vs not married), age at delivery, IQ, education, and cohort
Bashash *et al.* (2018)	Cohort (prospective) Mexico City/Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) participants [210]	Maternal urine during pregnancy Mean 0.85 (95% CI: 0.81, 0.90) mg/L	Children (ages 6–12 years)	ADHD: Conners' Rating Scales-Revised (CRS-R)	Significant associations between maternal urinary fluoride and CRS-R scores including Cognitive Problems + Inattention Index (adjusted β = 2.54; 95% CI: 0.44, 4.63), DSM-IV Inattention Index (adjusted β = 2.84; 95% CI: 0.84, 4.84), DSM-IV ADHD Total Index (adjusted β = 2.38; 95% CI: 0.42, 4.34), and ADHD Index (adjusted β = 2.47; 95% CI: 0.43, 4.50) Adjusted for gestational age, birth weight, child's gender, parity, age at outcome measurement, and maternal characteristics including smoking history (ever smoked vs. nonsmoker), marital status (married vs. not married), education, socioeconomic status, and cohort

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Canada					
Barberio *et al.* (2017b)	Cross-sectional General population/ Canadian Health Measures Survey (Cycles 2 and 3) [2,221]	Children's urine Mean Cycle 2: 32.06 (95% CI: 29.65, 34.46) µmol/L Mean Cycle 3: 26.17 (95% CI: 22.57, 29.76) µmol/L	Children (ages 3–12 years)	Learning disability, ADHD (Cycle 2 only): Parent or child self-report	Significant increase in adjusted OR for learning disability (adjusted OR = 1.02; 95% CI: 1.00, 1.03) only when Cycle 2 and 3 were combined using unadjusted urinary fluoride (associations no longer significant once adjusted for creatinine and specific gravity); no significant associations found between urinary fluoride and ADHD (only evaluated in Cycle 2) Adjusted for child's age, child's gender, household income adequacy, and highest attained education in the household
Riddell *et al.* (2019)	Cross-sectional General population/ Canadian Health Measures Survey (Cycles 2 and 3) [3,745]	Children's urine Mean (SD): 0.61 (0.39) mg/L; non-fluoridated water-0.46 (0.32) mg/L, fluoridated water-0.82 (0.54) Drinking water Mean (SD): 0.23 (0.24) mg/L; non-fluoridated water-0.04 (0.06) mg/L, fluoridated water-0.49 (0.22)	Children (ages 6–17 years)	Hyperactivity/inattention: Strengths and Difficulties Questionnaire (SDQ) ADHD: parent or self-reported physician diagnosis	Significantly increased risk of ADHD with fluoride in tap water (adjusted OR = 6.10 per 1-mg/L increase; 95% CI: 1.60, 22.8) or community water fluoridation status (1.21; 95% CI: 1.03, 1.42), but not with urinary fluoride; similar results observed with attention symptoms based on the SDQ scores Adjusted for child's age, child's gender, child's BMI, ethnicity, parental education, household income, blood lead, and smoking in the home

Table 6. Studies on Neurodevelopmental and Cognitive Function in Humans[a,b]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurological Outcome Summary
Adult Studies					
Jacqmin *et al.* (1994)	Cross-sectional France (Gironde and Dordogne)/elderly adults [3,490]	Drinking water Range: 0.03–2.03 mg	Adults (ages ≥ 65 years)	Cognitive function: Mini-Mental State (MMS) Examination	No significant increase in the prevalence of cognitive impairment with increasing fluorine quartiles No statistical adjustment for confounders
Li *et al.* (2016)	Cross-sectional China (Inner Mongolia)/adults [511]	Drinking water intake and urinary fluoride Mean (SD) levels reported for a subset of subjects with normal scores (2.23 [2.23] mg and 1.46 [1.04] mg/L, respectively) and subjects with cognitive impairment (3.62 [6.71] mg and 2.47 [2.88] mg/L, respectively)	Adults (ages ≥ 60 years)	Cognitive function: MMS Examination	Results suggested that degree of fluoride exposure was consistent with severity of skeletal fluorosis, and fluoride exposure may be a risk factor for cognitive impairment; however, neither water fluoride intake (adjusted ORs = 0.94 [95% CI: 0.85, 1.04] and 0.86 [95% CI: 0.69, 1.06] in the moderate and severe cognitive impairment groups, respectively) nor urinary fluoride levels (adjusted ORs = 1.12 [95% CI: 0.89, 1.42] and 1.25 [95% CI: 0.87, 1.81] in the moderate and severe cognitive impairment groups, respectively) were significantly correlated with cognitive impairment Adjusted for sex, age, education, marital status (married vs. not married), alcohol consumption (non-drinkers, light drinkers, moderate to heavy drinkers), smoking history (never smoker, ex-smoker, light smoker, heavy smoker), and serum homocysteine levels

*Three additional publications based on subsample (i.e., 50–60 children) of the larger Yu *et al.* (2018) cohort were identified (Zhao *et al.* 2020, Zhou *et al.* 2019, Zhao *et al.* 2019); however, these publications focused on mechanistic considerations and are not included in the study totals for IQ because the main study by Yu *et al.* (2018) is considered a better representation of the IQ results.

[a]Includes lower risk-of-bias studies.

[b]Definitions: **ADHD**: attention-deficit/hyperactivity disorder; **GCI**: General Cognitive Index; **GM**: geometric mean; **HOME**: Home Observation Measurement of the Environment; **IQ**: intelligence quotient; **MSCA**: McCarthy Scales of Children's Abilities; **WASI**: Wechsler Abbreviated Scale of Intelligence (Spanish version); **WISC-IV**: Wechsler Intelligence Scale for Children-Revised; **WRAML**: Wide Range Assessment of Memory and Learning; **WRAVMA**: Wide Range Assessment of Visual Motor Ability.

Overall Risk-of-bias Discussion of the Body of Evidence

The confidence rating for the body of evidence in humans was based on studies with the lowest potential for bias (i.e., studies rated probably low or definitely low risk of bias for at least two of the three key risk-of-bias questions). Each of these 28 studies (including 26 studies in children and 2 in adults) had little or no risk-of-bias concerns, and confidence in the body of evidence was not

downgraded for risk of bias. However, the remaining studies in the human body of evidence were rated as probably high or definitely high risk of bias for at least two key risk-of-bias questions or had other major concerns. Risk-of-bias ratings for individual studies for all questions are available in **Figure A3-1** and **Figure A3-3**. Among the studies with lower potential for bias (see **Figure A3-1** and **Figure A3-2**), the key risk-of-bias question with the most potential for bias was the potential for confounding. Potential confounding was a concern for 5 of the 26 lower risk-of-bias studies in children (see *Confounding* for further discussion). Among the studies with higher overall potential for bias, there were a number of risk-of-bias concerns, including potential confounding, poor exposure characterization, poor outcome assessment, and, in many cases, potential concern with participant selection (see **Figure A3-3** and **Figure A3-4**). Many of the studies (n = 32) included in the entire human body of evidence were initially published in a foreign language (mainly Chinese) and were either translated and published in volume 41 of the journal *Fluoride* (n = 19) or were translated by the Fluoride Action Network (n = 13) (http://fluoridealert.org/researchers/translations/complete_archive/). Most of these studies were considered to have high potential for bias due to lack of information across many key risk-of-bias questions. Therefore, in order to assess if the lack of information relevant to key risk-of-bias concerns was the result of a loss in translation, the original Chinese publications and the translated versions of the five studies that had the most potential for being included in the lower risk-of-bias group of studies were reviewed to determine if any of the risk-of-bias concerns could be addressed (An *et al.* 1992, Chen *et al.* 1991 [translated in Chen *et al.* 2008], Du *et al.* 1992 [translated in Du *et al.* 2008], Guo *et al.* 1991 [translated in Guo *et al.* 2008a], Li *et al.* 2009). For all five studies, the translations were determined to be accurate and there was no impact on the key risk-of-bias concerns.

Confounding

Potential confounding variables and/or effect modifiers that were considered key for all studies, populations, and outcomes included child's age, child's sex, and socioeconomic status (e.g., maternal education, household income, marital status, crowding). Additional potential confounding variables and/or effect modifiers considered important for this evaluation depending on the study population and outcome included race/ethnicity; maternal demographics (e.g., maternal age, body mass index [BMI]); parental behavioral and mental health disorders (e.g., ADHD, depression); smoking (e.g., maternal smoking status, secondhand tobacco smoke exposure); reproductive factors (e.g., parity); nutrition (e.g., BMI, growth, anemia); iodine deficiency/excess; minerals and other chemicals in water associated with neurotoxicity (e.g., arsenic, lead); maternal and paternal IQ; and quantity and quality of caregiving environment (e.g., Home Observation Measurement of the Environment [HOME] score). To be assigned a rating of probably low risk of bias for the key risk-of-bias question regarding confounding, studies were not required to address every potential confounder listed; however, studies were required to address the three key covariates for all studies, the potential for co-exposures if applicable (e.g., arsenic and lead, both of which could affect cognitive function), and any other potential confounders considered important for the specific study population and outcome. For example, studies of populations in China, India, and Mexico, where there is concern for exposures to high fluoride and high arsenic, were required to address arsenic, and smoking needed to be addressed in studies of adults when dementia was evaluated. In order to identify areas of China, India, and Mexico where arsenic is a concern, groundwater quality maps were evaluated (https://www.gapmaps.org/Home/Public#) (Podgorski and Berg 2020). If no arsenic measurements were available for the area, the arsenic groundwater quality predictions from the global arsenic 2020 map were used (Podgorski and Berg 2020). If an area had less than 50% probability of having arsenic levels greater than 10 µg/L (the WHO guideline concentration), the area was considered not to have an issue with arsenic that needed to be addressed by the study authors.

Among studies with lower risk-of-bias concerns, 21 of the 26 studies were considered to have lower potential for bias due to confounding. Relative to confounders considered key for all studies and populations (i.e., age, sex, and SES), one study did not address age, one study did not address sex, and two studies did not account for indicators of SES (e.g., parental education, household income). Nine of the 26 lower risk-of-bias studies accounted for maternal or family member smoking. Potential confounding related to co-exposure to arsenic was not accounted for in five lower risk-of-bias studies and was the main potential concern in these studies; however, three of these studies (Xiang *et al.* 2011, Wang *et al.* 2012, Xiang *et al.* 2003a) were still considered low risk of bias for confounding due to the fact that arsenic was observed in the low fluoride areas (which would bias the effect toward the null), but an effect was still observed. The other two studies did not address arsenic and were in areas that had potential for arsenic exposure to occur (Soto-Barreras *et al.* 2019, Valdez Jimenez *et al.* 2017). Seven studies did not consider co-exposures to lead; however, for all of these studies this co-exposure was considered unlikely to have an impact in these study populations as there was no evidence that lead was prevalent or occurring in relation to fluoride.

Although there is variability in the potential confounders considered and differences in populations evaluated, the consistency of the results among the lower risk-of-bias studies indicates that confounding is not a major concern in this body of evidence. Even though 5 of 26 lower risk-of-bias studies in children are considered to have higher potential for bias due to confounding that could not be ruled out for that specific population and outcome (see **Figure 6**), results were consistent across multiple populations; all but two of the lower risk-of-bias studies in children reported an association between higher fluoride exposure and lower IQ or another cognitive effect. Seven of the lower risk-of-bias studies confirmed the robustness of the results by conducting sensitivity analyses (Bashash *et al.* 2018, Bashash *et al.* 2017, Green *et al.* 2019, Yu *et al.* 2018, Wang *et al.* 2020a, Wang *et al.* 2020b, Till *et al.* 2020). None of the sensitivity analyses adjusting for additional confounders found meaningful shifts in the association between fluoride exposure and IQ or other measures of cognitive function. Bashash *et al.* (2017) found that adjusting for HOME score increased the association between maternal urinary fluoride and children's IQ. Bashash *et al.* (2018) examined several potential confounders in sensitivity analyses involving subsets of participants, including HOME scores, child contemporaneous fluoride exposure measured by child urinary fluoride adjusted for specific gravity, and maternal lead and mercury exposures. The authors reported that no sensitivity analyses indicated appreciable changes in the fluoride-related association with behaviors related to ADHD, nor did they find evidence of effect modification between sex and maternal urinary fluoride. Green *et al.* (2019) found that adjusting for lead, mercury, manganese, perfluorooctanoic acid, and arsenic concentrations did not substantially alter the associations with IQ. Sensitivity analyses by Yu *et al.* (2018) that adjusted for covariates (including age, sex, and socioeconomic status) did not find differences in the results compared to the primary analyses. Both Wang *et al.* (2020a) and Wang *et al.* (2020b) found the results of the sensitivity analysis to be the same as the results from the preliminary analysis. Till *et al.* (2020) found that adjusting for maternal urinary fluoride levels had little effect on the results.

As previously mentioned, most of the higher risk-of-bias studies in the human body of evidence did not address the potential confounders of greatest concern. Many of these studies conducted only simple statistical analyses without accounting for any potential confounders (50 of 61 higher risk-of-bias studies), and many studies did not report whether the study subjects were from areas of similar socioeconomic status or environmental conditions (n = 20 higher risk-of-bias studies). Potential confounding related to important co-exposures (e.g., arsenic and lead) was often not addressed in higher risk-of-bias studies. In studies where there was high exposure to fluoride via drinking water with high naturally-occurring fluoride or from the use of coal-containing fluoride, most researchers did not

account for potential exposures to arsenic, which is commonly found in coal and drinking water in fluoride-endemic areas of China and Mexico. In general, researchers did not account for potential exposures to lead; however, studies reporting lead levels in fluoride-endemic areas, including areas in China, often reported low levels of lead (Xiang *et al.* 2011, Choi *et al.* 2012, Seraj *et al.* 2012, Choi *et al.* 2015, Yu *et al.* 2018, Saxena *et al.* 2012, Xiang *et al.* 2003b). Therefore, lead is not assumed to be a common exposure in fluoride-endemic areas. Most of the studies did not account for smoking or socioeconomic status, nor did they provide information to lessen the risk-of-bias concern (e.g., list of study characteristics indicating no significant differences between comparison groups). However, as noted for the lower risk-of-bias studies, given the consistency of the evidence, confounding among higher risk-of-bias studies is likely less of a concern for the body of evidence as a whole than for any individual study.

Figure 6. Potential Confounders Considered in Lower Risk-of-bias Studies Conducted in Children

Study (Location)[1]	Potential Confounding Factors Considered[2]																Notes	Reported Effect of Fluoride[4]
	Subject Characteristics				Other Exposures					Socioeconomic Factors		Parental Characteristics				Other[3]		
	Age	Sex	Race/Ethnicity	Health Factors[3]	Arsenic	Smoking	Iodine	Lead	Other[3]	SES	Caregiving Environment (e.g., HOME score)	Demographics[3]	Reproductive Factors[3]	Health Factors[3]	IQ			
Overall RoB Rating for Confounding: Probably Low																		
Barberio 2017b (Canada)	√	√	–	–	√	–	–	√	–	√	–	–	–	–	–	–	–	Yes
Bashash 2017 (Mexico)	√	√	–	–	√	√	–	√	√	√	√	√	√	–	√	√	Other exposures: Hg, Ca Demographics: maternal age Reproductive: parity, birth order, birth weight, gestational age at delivery Other: cohort	Yes
Bashash 2018 (Mexico)	√	√	–	–	√	√	–	√	√	√	√	√	√	–	–	√	Other exposures: Hg, Ca Demographics: maternal age Reproductive: parity, birth order, birth weight, gestational age at delivery Other: cohort	Yes
Choi 2015 (China)	√	√	–	√	√	–	–	√	–	√	–	√	√	√	–	√	Health: subject Fe deficiency, illnesses before age 3, medical history of subject and caretakers Demographics: parental age Reproductive: parity Other: residential history	Yes
Cui 2018 (China)	√	√	√	√	√	√	√	–	–	√	–	√	√	√	–	√	Health: subject BMI, stress/anger/anxiety/depression, psychological trauma, having a cold, in relatives: thyroid diseases, cancer, mental retardation Demographics: maternal age Reproductive: abnormal birth Other: alcohol consumption, proximity to factory, physical activity, various dietary factors, environmental noise	Yes
Green 2019 (Canada)	√	√	√	–	√	√	–	√	√	√	√	√	√	–	–	√	Other exposures: Hg, Mn, PFOA Demographics: parental age, pre-pregnancy BMI Reproductive: parity, weeks of gestation, birth weight, maternal chronic condition during pregnancy Other: alcohol consumption, birth country, voiding interval in urine sampling, breastfeeding duration	Yes
Li et al., 2004 (translated in Li 2008a) (China)	√	√	–	–	√	–	–	–	–	√	–	√	√	–	–	–	Demographics: living habits, cultural background Reproductive: gestational age, birth weight, birth method Other: nutritional conditions	Yes
Riddell 2019 (Canada)	√	√	√	√	√	√	–	√	–	√	–	–	–	–	–	–	Health: subject BMI	Yes
Rocha-Amador 2007 (Mexico)	√	√	–	√	√	–	–	√	–	√	–	–	–	–	–	–	Health: subject height and weight by age, transferrin saturation	Yes
Saxena 2012 (India)	√	√	–	√	√	–	√	√	–	√	–	–	–	–	–	√	Health: subject height for age ratio, weight for height ratio, medical history Other: residential history	Yes
Seraj 2012 (Iran)	√	√	–	–	√	–	√	√	–	√	–	–	–	–	–	√	Other: fluorosis intensity	Yes
Sudhir 2009 (India)	√	√	–	–	√	–	–	√	–	√	–	–	–	–	–	√	Other: staple food consumed, liquids routinely consumed, aids used for oral hygiene maintenance (fluoridated or nonfluoridated)	Yes
Till 2020 (Canada)	√	√	√	–	√	√	–	–	–	√	√	–	–	–	–	√	Other: city	Yes
Trivedi 2012b (India)	√	√	–	–	√	–	√	–	–	√	–	–	–	–	–	–		Yes
Wang 2020a (China)	√	√	–	√	√	–	–	√	√	√	–	√	–	–	–	–	Other exposures: cadmium and mercury Health: subject BMI Demographics: mother and father migration, living habits Other: diet, industrial pollution within 1 km of living environment	Yes
Wang 2020b (China)	√	√	–	√	√	√	√	√	–	√	–	–	√	–	–	√	Health: subject BMI, exclusions based on diseases affecting intellegence, history of trauma or neurological disorders, positive screening test history Reproductive: low birth weight Other: alcohol consumption	Yes
Wang 2012 (China)	√	√	–	√	–	–	√	√	–	√	–	–	–	√	–	√	Health: medical conditions Other: transportation, natural environment, and lifestyle	Yes
Xiang 2003 (China)	√	√	–	–	–	–	√	√	–	√	–	–	–	–	–	–		Yes
Xiang 2011 (China)	√	√	–	–	–	–	√	√	–	√	–	–	–	–	–	–		Yes
Yu 2018 (China)	√	√	–	√	√	√	√	√	√	√	–	–	√	–	–	√	Health: subject BMI Reproductive: disease history during pregnancy, delivery conditions Other: dental fluorosis prevalence, consanguineous marriage	Yes
Zhang 2015b (China)	√	√	–	√	√	–	√	√	√	√	–	–	–	–	–	√	Health: subject physical and mental health status Other: thyroid hormone levels, residential history, having knowledge of fluorosis, COMT genotype	Yes

Study (Location)[1]	Potential Confounding Factors Considered[2]																Notes	Reported Effect of Fluoride[4]
	Subject Characteristics				Other Exposures					Socioeconomic Factors		Parental Characteristics				Other[3]		
	Age	Sex	Race/Ethnicity	Health Factors[3]	Arsenic	Smoking	Iodine	Lead	Other[3]	SES	Caregiving Environment (e.g., HOME score)	Demographics[3]	Reproductive Factors[3]	Health Factors[3]	IQ			
Overall RoB Rating for Confounding: Probably High																		
Cui 2020 (China)	–	√	–	√	√	√	√	–	–	√	–	√	√	√	–	√	Health: stress/anger/anxiety, having a cold, in relatives: mental retardation Demographics: maternal age Reproductive: abnormal birth, smoking and drinking during pregnancy Other: thyroid horomone levels	No
Ding 2011 (China)	√	–	–	–	√	–	√	√	–	–	–	–	–	–	–	–	–	Yes
Rocha-Amador 2009 (Mexico)	√	√	–	√	√	–	–	√	–	–	–	–	–	–	–	–	Health: subject height and weight by age	Yes
Soto-Barreras 2019 (Mexico)	√	√	–	–	–	–	–	–	–	√	–	–	–	–	–	–		No
Valdez Jimenez 2017 (Mexico)	√	√	–	–	–	–	–	–	–	√	–	√	√	√	–	√	Demographics: maternal age Health: pre-pregnancy history of drugs, vaccines, diseases Reproductive: prenatal history, parity, type of birth, week of birth, weight and length at birth, gestational age, Apgar and health conditions of the baby during the first month of life Other: infant feeding type (breastfeeding, formula)	Yes

Notes:

[1]Includes all lower risk-of-bias studies in children. Studies are organized as those with an overall risk-of-bias rating for confounding as probably low (green) followed by those with an overall risk-of-bias rating for confounding as probably high (yellow).

[2]Potential confounding factors and/or effect modifiers represented here are those considered important for this evaluation. See study details provided in HAWC for information on additional confounders.

Factors outlined in blue (subject age, subject sex, arsenic, SES) are considered key confounders.

A √ indicates that a factor was considered (and may or may not have been adjusted for in final model). For 'Other Exposures', a √ might also be used when a co-exposure was not expected to be an issue because there is no evidence to indicate that the co-exposure was prevalent or occurring in relation to fluoride. See risk-of-bias explanations in HAWC for details. A hyphen (–) indicates that the factor was not considered.

[3]See the "Notes" column for additional details.

[4]Extent of reported effects varies by study. "Yes" indicates that study authors reported one or more significant effects on IQ or other cognitive functions associated with fluoride exposure.

Exposure assessment

Exposure was assessed using a variety of methods in the human body of evidence. Studies provided varying levels of details on the methods used and employed different exposure characterization methods to group study subjects into exposed and reference groups. Exposure metrics included spot urine (from children or mothers during at least one trimester), serum, individual drinking water, intake from infant formula, estimated total exposure dose, municipal drinking water (with residence information), area of residence (endemic versus a non-endemic fluorosis area with or without individual validation of exposure), burning coal (with or without fluoride), and occupation type. Analytical methods to measure fluoride in biological or water samples also varied, some of which included atomic absorption, ion selective electrode methods, colorimetric methods, or the hexamethyldisiloxane microdiffusion method. Individual-level measures of exposure were generally considered more accurate than group-level measures; however, using group-level measures (e.g., endemic versus non-endemic area) in an analysis was less of a concern if the study provided water or urine fluoride levels from some individuals to verify that there were differences in the fluoride exposure between groups. Studies that provided results by area but also reported individual urinary or serum fluoride concentrations or other biochemical measures, including dental fluorosis in the children or urinary levels in mothers during pregnancy, were considered to have probably low risk of bias.

In general, there were few or no risk-of-bias concerns regarding exposure assessment in the lower risk-of-bias studies. Many of the lower risk-of-bias studies used individual urine or water measures with appropriate analyses. Urinary fluoride levels include all ingested fluoride and are considered a valid measure to estimate fluoride exposure (Villa *et al.* 2010, Watanabe *et al.* 1995); however, some concerns exist. Urinary fluoride is thought to reflect recent exposure but can be influenced by the timing of exposure (e.g., when water was last consumed, when teeth were last brushed). When compared to 24-hour urine samples, spot urine samples are more prone to these influences and can also be affected by differences in dilution; however, many studies attempted to account for dilution either using urinary creatinine or specific gravity. Good correlations between 24-hour samples and urinary fluoride concentrations from spot samples adjusted for urinary dilution have been described (Zohouri *et al.* 2006). Despite potential issues with spot urine samples, if authors made appropriate efforts to reduce the concern for bias, studies that used this metric were generally considered to have probably low risk of bias for exposure.

Although there are concerns related to using maternal urine samples, many studies provide evidence to suggest that urinary fluoride is a reasonable measure of exposure. Using three methods to account for urine dilution, Till *et al.* (2018) reported that adjusted risk estimates did not differ from unadjusted estimates. Analyzing the same study population as Till *et al.* (2018), Green *et al.* (2019) found that adjusting for time of urine collection or time of collection since last void during pregnancy did not substantially affect associations with IQ results in either boys or girls. In addition, adjusting the maternal urinary fluoride for creatinine did not substantially alter the association observed (Green *et al.* 2019). To provide a more accurate and sensitive measurement of maternal urinary fluoride than a single measurement provides, Green *et al.* (2019) only included participants with valid fluoride measurements at each trimester in their analysis. Several other studies also measured urinary fluoride multiple times throughout pregnancy (Bashash *et al.* 2018, Bashash *et al.* 2017, Green *et al.* 2019, Valdez Jimenez *et al.* 2017). Other studies demonstrated correlations between the urinary fluoride and fluoride in the drinking water, fluorosis, or estimated dose based on water (Green *et al.* 2019, Saxena *et al.* 2012, Zhang *et al.* 2015b, Ding *et al.* 2011, Choi *et al.* 2015, Yu *et al.* 2018). Till *et al.* (2018) demonstrated that there was a linear association between urinary fluoride concentrations in pregnant women and drinking water fluoride concentrations regardless of method to correct for urine dilution or whether or not

adjustments were made for dilution. Bashash *et al.* (2017) excluded exposure outliers but found that doing so did not change the results in a meaningful way. Taken together, these studies suggest that urinary fluoride is a reasonable measure of exposure despite some of the potential issues.

A frequent critical limitation among the higher risk-of-bias studies was lack of information regarding exposure or poor exposure characterization. Many of the higher risk-of-bias studies only compared subjects living in two regions with differing levels of fluoride exposure, and while most of them did provide some differentiation in levels of fluoride between the areas, limited or no individual exposure information was reported. Among studies that provided drinking water levels of fluoride in two areas being compared, sufficient information to determine if the individual study subjects were exposed to these levels was often not reported. Some studies also lacked information on fluoride analysis methods and timing of the exposure measurements. In some cases (n = 3), study areas that were considered endemic for dental and/or skeletal fluorosis were compared to non-endemic areas, or high-fluoride areas were compared to low-fluoride areas, with no other information provided on fluoride levels in the areas (Sun *et al.* 1991, Li *et al.* 2003 [translated in Li *et al.* 2008c], Ren *et al.* 1989 [translated in Ren *et al.* 2008]). While living in an area endemic for fluorosis could be an indicator of exposure, these studies did not specify if the study subjects themselves had fluorosis. Another study used only dental fluorosis as a measure of fluoride exposure in subjects that were all from an endemic area with similar drinking water fluoride levels (Li *et al.* 2010).

Outcome assessment

Studies included in this evaluation used a wide variety of methods to measure IQ and other cognitive effects. Measures of IQ were generally standardized tests of IQ; however, for these standardized methods to be considered low potential for bias they needed to be conducted in the appropriate population or modified for the study population. Because results of these tests can be subjective, it was important that the outcome assessors were blind to the fluoride exposure when evaluating the results of the tests. If the study reported that the assessor was blind to the exposure, this was assumed to mean that the outcome assessor did not have any knowledge of the exposure, including whether the study subjects were from high-fluoride communities.

The lower risk-of-bias studies have few concerns regarding outcome assessment. Four studies (Barberio *et al.* 2017b, Riddell *et al.* 2019, Sudhir *et al.* 2009, Wang *et al.* 2020a) had concerns for potential bias in the outcome assessment and that was due to either the use of self-reported of outcomes or the lack of accounting for blinding at the time of the outcome assessment in cases where there was potential concern. The remainder of the studies used appropriate measures of IQ or other cognitive effects for the study population. Seventeen of the studies reported blinding of the outcome assessors or correspondence with the study authors indicated that it was not likely an issue. For the remainder of the studies, it was assumed that the outcome assessors were most likely blind because exposure was assessed via urine or drinking water obtained at the same time as the outcome assessment.

Among the studies with higher risk of bias, the main limitation in the outcome assessment was the lack of reporting on whether the outcome was assessed without knowledge of exposure. Although there is little concern that the children's knowledge of their own exposure would bias the way they took the IQ tests, there is potential for bias if the tests were administered by an interviewer, or if the scoring of results could be subjective (e.g., drawing tests), and the interviewer or scorer had knowledge of the children's exposure. Most of the studies did not provide sufficient information on the person scoring or administering the tests or other information on the assessment methods to alleviate concerns for potential interviewer or reviewer bias. In some cases, the outcomes were not considered sensitive

measures (e.g., Seguin Form Board Test to test for IQ), or the test was not considered appropriate for the study population (e.g., a test validated in a western population was used on a rural Chinese population).

IQ in Children

The results from 17 studies (3 prospective cohort and 14 cross-sectional studies from 13 different study populations) with lower potential for bias that evaluated IQ in children (Bashash *et al.* 2017, Choi *et al.* 2015, Ding *et al.* 2011, Rocha-Amador *et al.* 2007, Saxena *et al.* 2012, Seraj *et al.* 2012, Xiang *et al.* 2003a, Xiang *et al.* 2011, Zhang *et al.* 2015b, Yu *et al.* 2018, Green *et al.* 2019, Cui *et al.* 2018, Wang *et al.* 2020b, Wang *et al.* 2012, Sudhir *et al.* 2009, Till *et al.* 2020, Trivedi *et al.* 2012) provide consistent evidence that exposure to fluoride is associated with lower IQ scores (see **Figure D1** through **Figure D7**); however, the analyses performed and the specific results varied by study. Consistent results between increased fluoride levels and lower IQ scores were seen across studies using different exposure measures [e.g., single serum samples (Xiang *et al.* 2011, Zhang *et al.* 2015b), single spot urine samples in children (Xiang *et al.* 2003a, Rocha-Amador *et al.* 2007, Ding *et al.* 2011, Saxena *et al.* 2012, Zhang *et al.* 2015b, Cui *et al.* 2018, Yu *et al.* 2018, Wang *et al.* 2020b), and prenatal maternal urinary measures (Bashash *et al.* 2017, Green *et al.* 2019)] (see **Figure D6** and **Figure D7**). The consistency also occurs across different study designs and study populations. There were two studies with lower potential for bias (Cui *et al.* 2020, Soto-Barreras *et al.* 2019) that did not provide evidence of an association between fluoride exposure and IQ, but evaluating the association between fluoride and IQ levels was not the primary focus in either of these studies.

The three prospective cohort studies all found an association between increasing fluoride exposure and lower IQ in children. Two of the studies (Green *et al.* 2019, Till *et al.* 2020) were based on the same study population but evaluated fluoride exposure differently. Bashash *et al.* (2017) observed a significant inverse association between children's IQ and maternal urinary fluoride during pregnancy (measured during all three trimesters and included if at least one measurement was available; an increase of 0.5 mg/L of maternal urinary fluoride was associated with a 2.5-point lower IQ score [95% CI: −4.12, −0.59]) in boys and girls combined (see **Figure D7**); however, the association between IQ level and children's urinary fluoride levels, while inverse, was not significant (single spot urine sample; an increase of 0.5 mg/L of child urinary fluoride was associated with a 0.89-point lower IQ score [95% CI: −2.63, 0.85]) (Bashash *et al.* 2017). Green *et al.* (2019) also observed a significantly lower IQ for boys associated with maternal urinary fluoride averaged across trimesters (a significant 4.49-point lower IQ score [95% CI: −8.38, −0.60] in IQ per 1-mg/L increase in maternal urinary fluoride); results were not significant in girls (2.40-point increase [95% CI: −2.53, 7.33] in IQ) or in boys and girls combined (1.95-point lower IQ score per 1-mg/L increase; 95% CI: −5.19, 1.28). Other measures of prenatal exposure (maternal fluoride intake or water fluoride concentrations) were associated with lower IQ scores in boys and girls combined although the authors did not report boys and girls separately, as they found no significant effect measure modification between child sex and fluoride exposure in these analyses (Green *et al.* 2019). Specifically, when evaluating the association between estimated maternal fluoride intake based on maternal water and beverage consumption during pregnancy and IQ in children, a 1-mg increase in daily maternal consumption of fluoride during pregnancy was associated with a significantly lower IQ score of 3.66 points in boys and girls combined (95% CI: −7.16, −0.15). Similarly, water fluoride concentrations for pregnant women from fluoridated areas (mean water fluoride levels of 0.59 ± 0.08 mg/L) versus pregnant women from non-fluoridated areas (mean water fluoride levels of 0.13 ± 0.06 mg/L) were associated with a significant 5.29-point lower IQ score per 1-mg/L increase in fluoride in both boys and girls combined (95% CI: −10.39, −0.19) (Green *et al.* 2019). Using the same study population as Green *et al.* (2019), but using fluoride intake from formula or water concentrations in

formula-fed versus breastfed infants, Till *et al.* (2020) observed a significantly lower performance IQ scores regardless of the comparison used. They did not observe any effect on verbal IQ, and full-scale IQ was only significantly lower in formula-fed infants using water fluoride concentrations as the exposure measure. All other comparisons showed negative associations but were not significant.

Cross-sectional studies also demonstrated a consistent association between fluoride and lower IQ scores. Rocha-Amador *et al.* (2007) observed significant negative correlations between IQ and both water and children's single spot urinary fluoride levels in a population in Mexico (adjusted β = −10.2 per log fluoride increase [CIs not reported] and −16.9 per log fluoride increase [CIs not reported], respectively) (see **Figure D7**). The authors also observed a significant inverse association between IQ and children's drinking water and single spot urinary arsenic levels (adjusted β = −6.15 [CIs not reported] and −5.72 [CIs not reported], respectively). Because fluoride and arsenic were highly correlated in the study area, the authors were not able to adjust for exposure to arsenic when evaluating the effects of fluoride exposure (Rocha-Amador *et al.* 2007). Ding *et al.* (2011) reported a negative dose-response relationship between children's single spot urinary fluoride levels and IQ (see **Figure D4**); after adjusting for age, using multiple linear regression, they found a 0.59-point lower IQ score (95% CI: −1.09, −0.08) per 1-mg/L increase in urinary fluoride (p-value < 0.0001) (see **Figure D7**). Cui *et al.* (2018) observed a significant association between log-transformed children's single spot urine fluoride and lower IQ scores (2.47-point lower IQ scores [95% CI: −4.93, −0.01] per unit increase in urinary fluoride), and the association was the strongest in subjects with the TT polymorphism in the dopamine receptor D2 gene which, according to the authors, probably results in a reduced D2 receptor density (12.31-point lower IQ score [95% CI: −18.69, −5.94] per unit increase in urinary fluoride) (Cui *et al.* 2018).

Although Green *et al.* (2019) observed a significant negative association between maternal fluoride levels and IQ scores in boys but not girls in a Canadian population, sex differences were not observed in a cross-sectional study conducted using children spot urine fluoride concentrations in China (Wang *et al.* 2020b). Wang *et al.* (2020b) evaluated boys and girls combined and separately and observed significant decreasing trends in the sexes both combined and alone by urinary fluoride quartiles. When evaluated as a continuous variable, spot urinary fluoride levels (per 1-mg/L increase) were significantly associated with lower IQ scores in both girls (−1.379 [95%CI: −2.628, −0.129]) and boys (−1.037 [95% CI: −2.040, −0.035]), as well as, combined (−1.214 [95%CI: −1.987, −0.442]). Green *et al.* (2019) did not find any sex difference when using water fluoride concentrations, but Wang *et al.* (2020b) found that based on water fluoride quartiles there was a significant trend in girls and in boys and girls combined but not in boys alone. While there was a decreasing trend in boys, the results did not achieve statistical significance (p = 0.077). However, when water fluoride levels were evaluated as a continuous variable (per 1 mg/L increase), there were significant associations between lower IQ scores in both girls (−1.649 [95%CI: −3.201, −0.097] and boys (−1.422 [95%CI: −2.792, −0.053], as well as, combined (−1.587 [95%CI: −2.607, −0.568]).

Other cross-sectional studies also observed consistent results across populations, but there were some slight variations based on exposure measurement, level of exposure, or based on the outcome measured. Choi *et al.* (2015) conducted a pilot study with 51 children in an area of China with a wide range of fluoride concentrations in the drinking water. Aside from observing no association between the square root block design test score and fluoride exposure from drinking water, the authors observed consistent negative associations between IQ measures and fluoride in children's single spot urine or drinking water and significant associations between specific tasks from an omnibus IQ test (i.e., significantly lower WISC-IV backward and total digit span scores) and fluoride exposure based on moderate or severe dental fluorosis in children (see **Figure D7**). While observing no association between

IQ and low children's single spot urinary fluoride levels (0.01–1.60 mg/L), Yu *et al.* (2018) observed significant negative associations (p values not reported) between IQ and median children's urinary fluoride levels (1.60–2.50 mg/L)—with an IQ score 2.67 points lower (95% CI: −4.67, −0.68) for every 0.5-mg/L increment of urinary fluoride—and high children's urinary fluoride levels at 2.50–5.54 mg/L with an IQ score of 0.84 points lower (95% CI: −2.18, 0.50) for every 0.5-mg/L increment of urinary fluoride (see **Figure D7**). The authors also reported a significant negative association between drinking water fluoride levels at 3.40–3.90 mg/L (4.29-point lower IQ score [95% CI: −8.09, −0.48] for every 0.5-mg/L increment of water fluoride); a 0.04-point lower IQ score (95% CI: −0.33, 0.24) was observed for 0.5-mg/L increments of water fluoride at levels of 0.20–3.40 mg/L. When comparing water fluoride concentrations of >1 mg/L to ≤1 mg/L, there was an increased risk (adjusted OR = 1.25; 95% CI: 0.69, 2.26) for marginal intelligence (i.e., IQ score = 70–79) and a decreased risk (adjusted OR = 0.47; 95% CI: 0.32, 0.71) of excellent intelligence (i.e., IQ score ≥ 130) (see **Figure D4**). Similar results were observed using children's urinary fluoride levels (adjusted OR for marginal intelligence = 1.44; 95% CI: 0.72, 2.91; adjusted OR for excellent intelligence = 0.49; 95% CI: 0.26, 0.93) (Yu *et al.* 2018).

Two lower risk-of-bias studies (Cui *et al.* 2020, Soto-Barreras *et al.* 2019) did not observe a significant association between fluoride and IQ in children; however, both studies only performed simple comparisons between IQ and fluoride exposure. Cui *et al.* (2020) studied children in the same region as Cui *et al.* (2018)—and possibly included some of the same subjects but over a longer timeframe—and did not observe a significant change in IQ score with increasing urinary fluoride levels. Although there was a 2-point drop in IQ between the lowest fluoride exposure group (i.e., spot urine fluoride < 1.6 mg/L) and the highest fluoride exposure group (i.e., ≥ 2.5 mg/L), the difference in IQ (112.16 ± 11.50 versus 110.00 ± 14.92) was not significant (p = 0.58). However, this study did not account for age in the analysis, even though they reported a significant difference in IQ score based on age (p < 0.001). Soto-Barreras *et al.* (2019) also did not find an association between various fluoride exposure metrics (water, spot urine, exposure dose, and fluorosis index) and IQ grade. However, this study only compared fluoride exposure levels within the 5 IQ grades and did not adjust for any potential confounders.

The results from 47 studies with higher potential for bias that evaluated IQ in children provide consistent supporting evidence of decrements in IQ associated with exposures to fluoride. Forty-one of the 47 studies reported an association between high fluoride exposure and lower IQ scores in children.

Meta-analysis

In response to the recommendations of the NASEM review of the September 6, 2019 draft monograph (NASEM 2020), a two-part meta-analysis was conducted. The first part was an update to two previous meta-analyses (Choi *et al.* 2012, Duan *et al.* 2018) of group-level exposures from studies that reported a comparison of the mean IQ score between two or more exposure groups. The second part was a new meta-analysis and included studies with more precise individual-level exposures (e.g., urine, water, fluoride intake). The meta-analysis protocol can be found with the revised systematic review protocol posted in September 2020 (https://ntp.niehs.nih.gov/go/785076).

Group-level exposures

For the group-level exposure meta-analysis, a comparison on the mean outcome measure (IQ score) was conducted across two exposure groups ("exposed" and "reference"). If there were more than two exposure groups, the highest exposure group was designated the exposed group and the lowest exposure group was designated the reference group. For studies that had more than one exposed group (n = 17), a sensitivity analysis was performed to evaluate the impact of using any exposed group compared to the reference group. Using mean IQ levels with measures of uncertainty (95% confidence

interval [CI], standard error [SE], and sample size [N]) for exposed versus reference groups, the standardized mean difference (SMD) and corresponding 95% CI were calculated for each study. Given the heterogeneity of the studies, random effects models were used to obtain the pooled effect estimate, calculated as a weighted SMD with a corresponding 95% CI. These methods are consistent with both the Choi *et al.* (2012) and Duan *et al.* (2018) meta-analyses. More detailed methods are provided in **Appendix 5** and in the meta-analysis protocol, which can be found with the revised systematic review protocol posted in September 2020 (https://ntp.niehs.nih.gov/go/785076).

Characteristics of the 46 studies that compared mean IQ scores between groups of children with different levels of fluoride exposure are shown in **Table A-1**. One study was conducted in New Zealand, 1 study was conducted in Mexico, 4 studies were conducted in Iran, 9 studies were conducted in India, and the remaining 31 studies were performed in China (**Table A-1**). Five study populations were exposed to fluoride from coal burning (Guo *et al.* 1991 [translated in Guo *et al.* 2008a], Li *et al.* 1994 [translated in Li *et al.* 2008b], Li *et al.* 1995, Li *et al.* 2009, Bai *et al.* 2014); otherwise, it is assumed that study populations were exposed to fluoride through drinking water. Measures of fluoride exposure included water fluoride (n = 28), dental fluorosis (n = 7), and other non-drinking water sources of exposure to fluoride (e.g., fluoride exposure from coal burning [n = 11]). Thirteen studies presented results for males and 12 studies reported results for females; 9 studies examined children < 10 years old and 11 studies examined children ≥ 10 years old. The CRT-RC was used to measure children's IQ in 23 studies. Other measures of IQ included Wechsler intelligence tests (Ren *et al.* 1989 [translated in Ren *et al.* 2008], Wang *et al.* 1996 [translated in Wang *et al.* 2008b], An *et al.* 1992, Broadbent *et al.* 2015), the Binet IQ test (Guo *et al.* 1991 [translated in Guo *et al.* 2008a], Xu *et al.* 1994), the Raven's test (Yao 1997, Yao *et al.* 1996, Seraj *et al.* 2006, Seraj *et al.* 2012, Eswar *et al.* 2011, Poureslami *et al.* 2011, Shivaprakash *et al.* 2011, Khan *et al.* 2015, Sebastian and Sunitha 2015, Mondal *et al.* 2016), the Raymond B Cattell test (Karimzade *et al.* 2014), the Japan IQ test (Sun *et al.* 1991, Zhang *et al.* 1998), the Index of Mental Capacity (Li *et al.* 1994 [translated in Li *et al.* 2008b]), the Sequin Form Board test (Nagarajappa *et al.* 2013), and other tests using a doctor-prepared questionnaire (Trivedi *et al.* 2012, Trivedi *et al.* 2007). This meta-analysis includes 27 studies that were also included in Choi *et al.* (2012) and 25 studies that were also included in Duan *et al.* (2018). Also included in this meta-analysis were an additional 3 studies published since the Duan *et al.* (2018) publication and 11 studies that were not captured in either of the previous meta-analyses. Overall, the updated group-level results were highly consistent with these previous meta-analyses (Choi *et al.* 2012, Duan *et al.* 2018) (**Table A5-1**).

The random-effects pooled SMD estimated from the 46 studies included in the meta-analysis was −0.50 (95% CI: −0.61, −0.39) (**Table A5-1**, **Figure A5-1)**. There was evidence of heterogeneity (I^2 = 89%, p < 0.001, **Table A5-1**) and publication bias (funnel plot and Egger's p < 0.001, Begg's p = 0.08; **Figure A5-2**, **Figure A5-3**). Eliminating publication bias through trim-and-fill analysis supports the results with an adjusted pooled effect estimate of – 0.42 (95% CI: −0.54, 0.30) (**Figure A5-4**). Among the 46 studies, all but two showed SMD estimates that indicated an inverse association, ranging from −5.34 (95% CI: −6.34, −4.34) to −0.04 (95% CI: −0.45, 0.36). The studies with a positive association (Broadbent *et al.* 2015) reported an SMD estimate of 0.01 (95% CI: −0.19, 0.22) to 0.13 (95% CI: −0.16, 0.42). Three studies (Aravind *et al.* 2016, Kundu *et al.* 2015, Razdan *et al.* 2017) were excluded from the main analysis due to uncertainties about the way the intelligence assessment for children was performed, but sensitivity analyses that included these studies did not reveal any substantial changes in the pooled SMD estimate (−0.57 [95% CI: −0.69, −0.45]) (see **Figure A-35**).

Several subgroup analyses, discussed in **Appendix 5** and outlined in the meta-analysis protocol (found with the revised systematic review protocol posted in September 2020

[https://ntp.niehs.nih.gov/go/785076]), included risk of bias, gender, age group, country, outcome assessment type, and exposure assessment type. Among the lower risk-of-bias studies (n = 9), the random-effects pooled SMD was −0.31 (95% CI: −0.52, −0.10) with an I^2 of 87% and heterogeneity test p-value < 0.001 (**Table A5-1** and **Figure A5-6**). There was no evidence of publication bias (funnel plot and Egger's p = 0.72, **Figure A5-7** and **Figure A5-8**). Among the higher risk-of-bias studies (n = 37), the random-effects pooled SMD was −0.56 (95% CI: −0.68, −0.43) with an I^2 of 88% and heterogeneity test p-value < 0.001 (**Table A5-1** and **Figure A5-6**). There was evidence of publication bias among the higher risk-of-bias studies (funnel plot and Egger's p < 0.001, **Figure A5-7** and **Figure A5-8**); eliminating publication bias through trim-and-fill analysis supports the results with an adjusted pooled SMD estimate of −0.35 (95% CI: −0.50, −0.21) (**Figure A5-7** and **Figure A5-9**).

Subgroup analyses by gender, age group, country, outcome assessment type, and exposure assessment type further support the consistent and robust pattern of results (**Table A5-1**). Except for the subgroup analysis of the four studies from Iran, heterogeneity remained at an I^2 of ≥ 70% when the analyses were restricted by subgroup. Sensitivity analyses that removed an outlier (Khan *et al.* 2015) or compared all exposed groups versus the reference (i.e., exposed groups were combined if a study reported more than one exposed group) also did not appreciably change the results (**Figure A-45**, **Table A5-1**, and **Figure A-25**).

Individual-level exposures

The individual-level exposure meta-analysis included 6 studies with individual-level exposures that reported effect estimates as beta coefficients and included a 95% CI or SE. Characteristics of the studies with individual-level exposures are shown on **Table B-1**. All studies included in this meta-analysis were considered lower risk of bias. Adjusted effect estimates were used, and if results from multiple models were reported within a single study, the most adjusted results were selected. (For more details, see the meta-analysis protocol, which can be found with the revised systematic review protocol posted in September 2020 [https://ntp.niehs.nih.gov/go/785076].) To ensure consistent units across studies, units of fluoride exposure were transformed to 1 mg/L. For Bashash *et al.* (2017), Yu *et al.* (2018), and Till *et al.* (2020), units of exposure were transformed from 0.5 mg/L to 1 mg/L. For Cui *et al.* (2018), units of exposure were transformed from 1 log mg/L to 1 mg/L. Cui *et al.* (2018) reported an association between IQ and log transformed exposure. A sensitivity analysis was performed to evaluate the impact of using Cui *et al.* (2018), since the relationship between IQ and exposure evaluated in this study was not linear (as was the case among the other studies included). Yu *et al.* (2018) reported estimates from piecewise linear regression models and provided three ranges for urinary fluoride exposure (low 0.01–1.60 mg/L, medium 1.60–2.50 mg/L, high 2.50–5.54 mg/L) and two ranges for water fluoride (low 0.20–3.40 mg/L and high 3.40–3.90 mg/L). Since these piecewise effect estimates are likely correlated, the study-specific pooled effect estimates were used for urine and water fluoride exposures for the overall effect meta-analysis. A sensitivity analysis was performed to evaluate the impact of using pooled estimates rather than piecewise estimates from Yu *et al.* (2018).

For studies with overlapping populations (i.e., multiple studies that used the same cohort), results were selected considering the following factors: most appropriate exposure metric, exposure range, exposure period, number of subjects, and statistical adjustment for potential confounders. In the overall effect analysis, for studies reporting multiple measures of fluoride exposure, the results associated with measured or estimated individual-level exposures, biomarker levels (such as urinary fluoride), or fluoride intake levels were prioritized over water fluoride concentrations (revised protocol posted in September 2020 [https://ntp.niehs.nih.gov/go/785076]); however, subgroup analyses by exposure metric (urinary fluoride, fluoride intake, and water fluoride) were also performed. Yu *et al.* (2018) and

Wang *et al.* (2020b) used the same study cohort of children recruited in 2015 from the rural areas of Tianjin City, China. Since Wang *et al.* (2020b) (n = 571) used a subset of the original study sample from Yu *et al.* (2018) (n = 2,668), only results from Yu *et al.* (2018) were included in the meta-analysis. A sensitivity analysis was performed to evaluate the impact of using the effect estimate from Wang *et al.* (2020b) rather than the pooled effect estimate from Yu *et al.* (2018). Green *et al.* (2019) and Till *et al.* (2020) used the same Maternal-Infant Research on Environmental Chemicals (MIREC) cohort that reported drinking tap water in 10 Canadian cities with the studies overlapping for 398 mother-child pairs. Both studies reported effect estimates for maternal urinary fluoride (MUF) and water fluoride concentrations. In the Green *et al.* (2019) study, 512 mother-child pairs had MUF data (and all covariates) compared to 398 pairs in Till *et al.* (2020). Water fluoride levels were available for 420 pairs in Green *et al.* (2019) compared to 398 pairs in Till *et al.* (2020). Both studies reported effect estimates adjusted for maternal education, maternal race, child's sex, HOME total score, and secondhand smoke status in the child's home. In addition, Till *et al.* (2020) adjusted for child's age at IQ testing (the age range for all children was 3–4 years old). Because of the larger sample size and covariate adjustments were similar, results from Green *et al.* (2019) were included in the main analysis. However, because of the more adjusted estimates from Till *et al.* (2020) compared to Green *et al.* (2019), a sensitivity analysis was performed using the water fluoride result for formula-fed children and the MUF result from Till *et al.* (2020). For fluoride from intake, the estimates from both studies were used since they represent total fluoride intake (Green *et al.* 2019) and infant fluoride intake from formula (Till *et al.* 2020).

The overall pooled effect estimate based on 6 lower risk-of-bias studies with individual-level measures of exposure shows that a 1-mg/L increase in urinary fluoride was associated with a statistically significant lower IQ score of 1.40 (95% CI: −2.33, −0.47) points. Studies with individual-level urinary fluoride measures had evidence of moderate heterogeneity (I^2 = 46%, p = 0.101; **Table A5-2**, **Figure A5-16**). Eliminating publication bias through trim-and-fill analysis supports the conclusion that a 1-mg/L increase in individual-level exposure to urinary fluoride was associated with lower IQ, with an adjusted pooled effect estimate of −0.82 (95% CI: −1.81, 0.17) (**Figure A5-19**). Fluoride intake and water fluoride were also significantly associated with an IQ score of 3.31 points lower (95% CI: −6.12, −0.50) and 4.77 points lower (95% CI: −9.10, −0.45), respectively (**Table A5-2**); however, the results for both metrics were based on two studies each and should be interpreted with caution.

No substantial changes in the pooled effect estimates were seen in sensitivity analyses to evaluate the following scenarios: using the piecewise estimates from Yu *et al.* (2018) (−1.37, 95% CI: −2.38, −0.37) (**B-1**); using effect estimates from Wang *et al.* (2020b) rather than Yu *et al.* (2018) (−1.24, 95% CI: −1.94, −0.54) (**Figure B-6**); and using the water fluoride result for formula-fed children and MUF result from Till *et al.* (2020) rather than effect estimates from Green *et al.* (2019) (−1.50, 95% CI: −2.44, −0.57) (**Figure B-11**).

Other Neurodevelopmental or Cognitive Effects in Children

Among the studies with lower potential for bias, the results from three prospective cohort studies (Bashash *et al.* 2017, Valdez Jimenez *et al.* 2017, Bashash *et al.* 2018) and six cross-sectional studies (Li *et al.* 2004 [translated in Li *et al.* 2008a], Rocha-Amador *et al.* 2009, Choi *et al.* 2015, Barberio *et al.* 2017b, Wang *et al.* 2020a, Riddell *et al.* 2019) based on seven study populations provide mostly consistent results for associations of fluoride exposure with cognitive impairment in children other than decrements in IQ, such as hand-eye coordination, neurobehavioral assessment, behavioral capacity, and learning disabilities (see **Figure D8** through **Figure D10**). Because IQ cannot be assessed in infants, other neurodevelopmental tests were conducted. Two studies (Li *et al.* 2004 [translated in Li *et al.* 2008a],

Valdez Jimenez *et al.* 2017), based in China and Mexico, evaluated neonates (within 3 days of birth) or infants (3–15 months) (see **Figure D8** and **Figure D10**).

In neonates, the high fluoride group (based on a single maternal urine fluoride level just prior to birth [3.58 ± 1.47 mg/L] compared to controls [1.74 ± 0.96 mg/L]) had significant lower ($p < 0.05$) total neurobehavioral assessment scores (38.28 ± 1.10 in controls compared to 36.48 ± 1.09 in high fluoride group) and total behavioral capacity scores (11.34 ± 0.56 in controls compared to 10.05 ± 0.94 in high fluoride group) as measured by a standard neonatal behavioral neurological assessment (NBNA) method (Li *et al.* 2004 [translated in Li *et al.* 2008a]). In infants, the Mental Development Index (MDI)—which measures functions including hand-eye coordination, manipulation, understanding of object relations, imitation and early language development—was significantly negatively correlated with maternal urinary fluoride in both the first and second trimesters (adjusted βs = −19.05 with standard error of 8.9 for first trimester and −19.34 with standard error of 7.46 for second trimester) (Valdez Jimenez *et al.* 2017). This study did not find an association between maternal fluoride during any trimester and Psychomotor Developmental Index (PDI), which measures gross motor development (adjusted βs = 6.28 and 5.33 for first and second trimesters, respectively; no variance provided) (Valdez Jimenez *et al.* 2017). The General Cognitive Index (GCI) of the McCarthy Scales of Children's Abilities (MSCA) in 4-year-old children was significantly negatively associated with maternal creatinine-adjusted urinary fluoride levels during pregnancy (collected during each trimester) even after adjusting for maternal bone lead (adjusted β = −3.15 [95% CI: −5.42, −0.87] in a model adjusting for main covariates (e.g., gestational age, weight at birth, sex, maternal smoking, and indicators of socioeconomic status); adjusted β = −5.63 [95% CI: −8.53, −2.72] in a model limited to a subset of cases who had data on maternal bone lead and adjusted for main covariates and maternal bone lead) (Bashash *et al.* 2017) (see **Figure D10**). Choi *et al.* (2015), however, evaluated cognitive function endpoints in addition to IQ and found no significant associations between concurrent water or urinary fluoride levels and Wide Range Assessment of Visual Motor Ability (WRAVMA) scores, finger tapping, and the grooved pegboard test although there were some significant associations based on degree of fluorosis (see **Figure D10**). Another study using construction and memory scores in children 6–11 years old observed statistically significant lower scores with increasing concurrent child single spot urinary fluoride even after adjusting for age ($p < 0.05$; −0.29 and −0.27 for copy and immediate recall, respectively [CIs not reported]); however, scores were not significantly associated with urinary arsenic levels (−0.05 and 0.02 for copy and immediate recall, respectively [CIs not reported]) (Rocha-Amador *et al.* 2009) (see **Figure D9**).

Barberio *et al.* (2017b) evaluated learning disabilities in children 3–12 years of age, including ADHD, attention deficit disorder (ADD), and dyslexia, as part of the Canadian Health Measures Survey and found a small but significantly increased risk in self-reported (children 12 years of age) or parent- or guardian-reported (children 3–11 years of age) learning disabilities associated with higher spot urinary fluoride levels in children (adjusted OR = 1.02; 95% CI: 1.00, 1.03) (see **Figure D11**); however, significant associations were not observed in analyses using creatinine- or specific gravity-adjusted urinary fluoride (Barberio *et al.* 2017b). Barberio *et al.* (2017b) also reported no associations between single spot urinary fluoride and ADHD in children ages 3 to 12 years. Riddell *et al.* (2019) used the same Canadian Health Measured Survey, but evaluated children 6–17 years old. Riddell *et al.* (2019) found a significantly increased risk for ADHD diagnosis with both tap water fluoride (adjusted OR per 1-mg/L increase = 6.10; 95% CI: 1.60, 22.8) and community water fluoridation status (adjusted OR = 1.21; 95% CI: 1.03, 1.42). A similar increase in hyperactivity-inattention symptoms score based on the Strengths and Difficulties Questionnaire was observed with both tap water fluoride (adjusted β per 1-mg/L increase = 0.31; 95% CI: 0.04, 0.58) and community fluoridation status (adjusted β = 0.11; 95% CI: 0.02, 0.20). As was observed with Barberio *et al.* (2017b), Riddell *et al.* (2019) did not observe an association with either

ADHD diagnosis (adjusted OR per 1-mg/L increase = 0.96; 95% CI: 0.63, 1.46) or hyperactivity-inattention symptoms (adjusted β = 0.31; 95% CI: −0.04, 0.66) and specific-gravity-adjusted spot urinary fluoride concentrations. Bashash *et al.* (2018) evaluated behaviors associated with ADHD in children ages 6–12 years using the Conners' Rating Scales-Revised (CRS-R) and observed significant associations between maternal urinary fluoride (measured during each trimester) and ADHD-like symptoms, particularly those related to inattention (an increase in 0.5 mg/L of maternal urinary fluoride was associated with a 2.84-point increase [95% CI: 0.84, 4.84] in the DSM-IV Inattention Index and a 2.54-point increase [95% CI: 0.44, 4.63] in the Cognitive Problems and Inattention Index). These two scales contributed to the global ADHD Index and the DSM-IV ADHD Total Index, which were also associated with higher levels of prenatal fluoride exposure (an increase of 0.5 mg/L in maternal urinary fluoride was associated with a 2.38-point increase [95% CI: 0.42, 4.34] in the DSM-IV ADHD Total Index and a 2.47-point increase [95% CI: 0.43, 4.50] in the ADHD Index) (see **Figure D10**). Significant associations were not observed between maternal urinary fluoride concentrations during pregnancy and child performance on measures of hyperactivity nor were there any significant results in children using the Connors' Continuous Performance Test (CPT-II, 2nd Edition), a computerized test of sustained attention and inhibitory control (Bashash *et al.* 2018). Wang *et al.* (2020a) also used a Connors' Parent Rating Scale (Chinese version), but only found a significant association between spot urinary fluoride concentrations (model adjusted for creatinine) and psychosomatic problems (adjusted OR for T-score > 70 = 1.97; 95% CI: 1.19, 3.27 and adjusted β = 4.01; 95% CI: 2.74, 5.28). No associations were found between spot urinary fluoride and ADHD index or other behavioral measures.

Higher risk-of-bias studies (n = 5) also provide some evidence of associations of fluoride exposure with neurodevelopmental or cognitive effects in children other than effects on IQ, but the results are inconsistent with heterogeneous outcomes (Li *et al.* 1994 [translated in Li *et al.* 2008b], Shannon *et al.* 1986, Malin and Till 2015, Morgan *et al.* 1998, Mustafa *et al.* 2018).

Cognitive Effects in Adults

Results from two lower risk-of-bias studies in adults did not find consistent evidence for an association between cognitive impairment (based on the Mini-Mental State Examination) and exposure to fluoride (Jacqmin *et al.* 1994, Li *et al.* 2016). Jacqmin *et al.* (1994) did not find an association between drinking water fluoride and cognitive impairment in populations in France (n = 3,490) and found prevalence rates of cognitive impairment to be the same regardless of fluoride exposure (see **Figure D12**). In an analysis of 38 cognitively-impaired cases and 38 controls matched for several confounders including age, gender, education, alcohol consumption, and smoking, Li *et al.* (2016) did find significantly higher urinary fluoride levels and skeletal fluorosis scores in the cognitively-impaired group compared with the control group; however, the authors found no significant correlation between cognitive impairment and total daily water fluoride intake (adjusted ORs = 0.94 [95% CI: 0.85, 1.04] and 0.86 [95% CI: 0.69, 1.06] in the moderate and severe cognitive impairment groups, respectively) or urinary fluoride levels (adjusted ORs = 1.12 [95% CI: 0.89, 1.42] and 1.25 [95% CI: 0.87, 1.81] in the moderate and severe cognitive impairment groups, respectively) in subjects from fluorosis-endemic areas of China (n = 511).

Higher risk-of-bias studies (n = 7) provide some evidence of cognitive impairment in adults associated with exposure to fluoride. In aluminum factory workers (exposed to gaseous and particular fluoride emissions during the production of aluminum metal), significant decreases in IQ (Duan *et al.* 1995), diminished performance on several neurobehavioral core battery tests (NCTBs) (Guo *et al.* 2001 [translated in Guo *et al.* 2008b]), and impaired psychomotor performance and memory were observed (Yazdi *et al.* 2011). One study conducted on adult subjects with fluorosis (dental and skeletal) from a fluorosis-endemic area compared with healthy subjects from a non-endemic area observed significant

differences for some cognitive function tests (i.e., tests of speech fluency, recognition, and working memory) but not others and generally did not observe a significant change in IQ except in the operation scores (Shao 2003). One prospective cohort study evaluated exposure to fluoride in children at age of 5 years based on whether or not the children resided in areas with community water fluoridation or used fluoride toothpaste or fluoride tablets, and found no clear differences in IQ scores of the subjects at age 38 years (Broadbent *et al.* 2015). One additional study suggested that populations living in areas with higher drinking water fluoride had lower levels of dementia (Still and Kelley 1980); however, the study was not focused on effects of fluoride, but rather if fluoride was able to reduce the risk associated with aluminum by competing with aluminum and reducing the aluminum bioavailability. Therefore, the study was considered inadequate to evaluate the effects of fluoride on dementia (Still and Kelley 1980). A more recent study in Scotland evaluated dementia rates associated with aluminum and fluoride drinking water concentrations and observed an increase in dementia only in the highest quartile of fluoride (56.3 µg/L) compared to the lowest quartile (<44.4 µg/L), but found a significant increase with all quartiles of aluminum compared with the reference group (Russ *et al.* 2019). In addition to studies that reported on cognitive impairment and exposure to fluoride, two studies were identified that reported effects on motor and sensory function (Rotton *et al.* 1982) and a higher prevalence of self-reported headaches, insomnia, and lethargy (Sharma *et al.* 2009).

Mechanistic Data in Humans

Eight lower risk-of-bias studies were available that evaluated mechanistic data in humans associated with fluoride exposure that was considered potentially relevant to neurological effects, including effects on thyroid hormones in children (Singh *et al.* 2014, Zhang *et al.* 2015b, Kumar *et al.* 2018), adults (Kheradpisheh *et al.* 2018b, Kheradpisheh *et al.* 2018a, Malin *et al.* 2018), or children and adults combined (Barberio *et al.* 2017a). In addition, some studies evaluated self-reported thyroid conditions in children and adults combined (Barberio *et al.* 2017a) and thyroid diseases in adults (Kheradpisheh *et al.* 2018b, Peckham *et al.* 2015) (see **Figure A3-5** and **Figure A3-6**). Although the lower risk-of-bias studies provide some evidence of mechanistic effects (primarily changes in thyroid stimulating hormone [TSH] levels in children), the studies were too heterogeneous or limited in number to make any determination on mechanism (see **Figure 7**).

Among the seven lower risk-of-bias studies that reported on changes in thyroid hormones, three studies were conducted in children (Zhang *et al.* 2015b, Singh *et al.* 2014, Kumar *et al.* 2018) and reported increases in TSH levels. Zhang *et al.* (2015b) reported significant increases in TSH in children from a fluorosis-endemic area (median fluoride drinking water concentration = 1.40 mg/L; interquartile range = 1.23–1.57 mg/L) compared with a non-fluorosis-endemic area (median fluoride drinking water concentration = 0.63 mg/L; interquartile range = 0.58–0.68 mg/L), while 3,5,3'-triiodothyronine (T_3) or thyroxine (T_4) were not significantly different between the two groups. Similarly, Singh *et al.* (2014) observed significantly higher TSH levels in children without dental fluorosis who lived in a fluorosis-endemic area (fluoride drinking water concentrations of 1.6–5.5 mg/L) compared with children without dental fluorosis who lived in a non-fluorosis-endemic area (fluoride drinking water concentrations of 0.98–1.00 mg/L). Higher TSH levels in children with dental fluorosis from the fluorosis-endemic area compared with children without dental fluorosis from the non-fluorosis-endemic area were observed but did not reach statistical significance. Significant differences in T_4 or T_3 were not observed between groups (Singh *et al.* 2014). Kumar *et al.* (2018) also observed a significant increase in TSH levels in children from a fluorosis endemic area (1.5–5.8 mg/L fluoride) compared with a control area (0.94–1.08 mg/L fluoride). There were also decreases in T_3 and T_4, but results were not statistically significant.

Barberio *et al.* (2017a) evaluated fluoride effects on TSH levels in children and adults combined and found no relationship between fluoride exposure (measures in urine and tap water) and TSH levels. In the one study that evaluated thyroid hormone levels in adults but not children, Kheradpisheh *et al.* (2018b) found a significant increase in TSH associated with higher fluoride concentrations in drinking water in both adults with and without thyroid diseases such as hypothyroidism, hyperthyroidism, thyroid nodules, or thyroid cancer. Significant increases in T_3 were associated with higher fluoride in drinking water in adults without thyroid diseases, but increases in T_3 were not significant in adults with thyroid diseases. A significant association between T_4 and higher fluoride in drinking water was not observed in adults with or without thyroid diseases (Kheradpisheh *et al.* 2018b).

Other than changes in hormone levels, there is limited evidence of fluoride-related mechanistic effects in the three lower risk-of-bias studies that evaluated thyroid-related effects. Barberio *et al.* (2017a) found no relationship between fluoride exposure and self-reported thyroid conditions in children and adults (children were older than 12). Kheradpisheh *et al.* (2018b) also found no association between fluoride exposure and hypothyroidism in an adult population in Iran. One study found a significantly higher prevalence of hypothyroidism in areas with higher fluoride concentrations in drinking water (>0.7 mg/L) compared with areas with lower fluoride drinking water concentrations (≤0.7 mg/L) (Peckham *et al.* 2015).

Several higher risk-of-bias studies were available that evaluated potential mechanistic data in humans associated with fluoride exposure, including effects on thyroid hormones mostly in children (n = 11 studies); catecholamines in adults (Michael *et al.* 1996) or in subjects of unknown ages (Chinoy and Narayana 1992); acetylcholinesterase (AChE) or serotonin levels in children (Singh *et al.* 2013, Lu *et al.* 2019); brain histopathology or biochemistry in aborted fetuses (Du *et al.* 1992 [translated in Du *et al.* 2008], Yu *et al.* 1996 [translated in Yu *et al.* 2008]); and mitochondrial fission/fusion molecules in children (Zhao *et al.* 2019). Similar to the lower risk-of-bias studies, the higher risk-of-bias studies provide some evidence of mechanistic effects (primarily changes in TSH levels in children); however, the data are insufficient to identify a clear mechanism by which fluoride causes neurodevelopmental or cognitive effects in humans.

Among higher risk-of-bias studies (see **Figure A3-7** and **Figure A3-8**), varying results were reported in 11 studies that evaluated fluoride exposure and effects on thyroid hormones, and a few of these studies (Lin *et al.* 1991, Yang *et al.* 1994 [translated in Yang *et al.* 2008], Wang *et al.* 2001) were complicated by high or low iodine in the high fluoride area. When considering fluoride effects on each of the hormones individually, similar to results from lower risk-of-bias studies, the most consistent evidence of fluoride-associated effects on a thyroid hormone was reported as changes in TSH levels in children, although there was some variation in the direction of effect. Six of the nine higher risk-of-bias studies that evaluated changes in TSH levels in children reported increases in TSH levels with higher fluoride (Lin *et al.* 1991, Susheela *et al.* 2005, Wang *et al.* 2001, Yang *et al.* 1994 [translated in Yang *et al.* 2008], Yao *et al.* 1996, Yasmin *et al.* 2013). Two of the nine higher risk-of-bias studies reported decreases in TSH levels in children with higher fluoride (Khandare *et al.* 2017, Khandare *et al.* 2018). One of the nine studies found no significant alterations in TSH levels in children from fluorosis-endemic areas (Hosur *et al.* 2012) (see **Figure 8**).

When considering fluoride-associated effects on TSH, T_3, and T_4 levels together, studies that evaluated changes in all three thyroid hormones reported varying combinations of increases, decreases, or no changes in levels across the three hormones, although among the eight lower and higher risk-of-bias studies that evaluated the effects of fluoride exposure on TSH, T_3, and T_4 levels and reported increases

in TSH levels in children, seven of the eight studies found no alterations in T_3 levels (one study found an increase in T_3), and six of the eight studies found no alterations in T_4 levels (two studies found an increase in T_4). Studies also displayed variation by age in fluoride-associated effects on TSH, T_3, and T_4. Due to the dynamic relationship between the thyroid gland, the pituitary gland, and the production and clearance of TSH, T_3, and T_4, the variations in results are not unexpected and do not eliminate the possibility of a mechanistic link between thyroid effects and neurodevelopmental or cognitive effects; however, the data do not support a clear indication that thyroid effects are a mechanism by which fluoride causes these effects in humans.

In addition to evaluating thyroid hormone levels, a few higher risk-of-bias studies evaluated other mechanistic data associated with fluoride exposure; however, the data are insufficient to identify a clear mechanism by which fluoride might cause neurodevelopmental or cognitive effects in humans. Serum epinephrine and norepinephrine were significantly increased in a fluoride-endemic region (not reported whether subjects were children or adults) compared to a non-endemic region (Chinoy and Narayana 1992). Serum adrenaline and noradrenaline were significantly increased in adults in a fluoride-endemic area (fluoride in the drinking water ranged from 1.0–6.53 ppm) compared to a control area (fluoride in the drinking water ranged from 0.56–0.72 ppm) (Michael *et al.* 1996). Serum AChE was significantly reduced in children from a high fluoride region compared to a lower fluoride region (Singh *et al.* 2013). Serum serotonin was significantly increased in children from Turkey who were drinking water containing 2.5 mg/L of fluoride compared to children drinking bottled water or water containing <0.5 mg/L of fluoride (Lu *et al.* 2019). Aborted fetuses from high fluoride areas in China were found to have histological changes in the brain and significant changes in neurotransmitter levels compared to a control area (Du *et al.* 1992 [translated in Du *et al.* 2008], Yu *et al.* 1996 [translated in Yu *et al.* 2008]).

There are also two more recent lower risk-of-bias studies that evaluated polymorphisms in dopamine-related genes; however, a determination on mechanism cannot be made at this time due to the limited number of studies. For children (10–12 years old) with a Val158Met polymorphism in the *COMT* gene (i.e., catechol-O-methyltransferase), which results in slower degradation and greater availability of dopamine within the brain, a stronger association between increasing urinary fluoride levels and decreasing IQ was reported (Zhang *et al.* 2015b). For children (7–12 years old) with a dopamine receptor-2 (DRD2) Taq 1A polymorphism (which is involved in reduced D2 receptor density and availability) and the TT (variant) genotype, a significant inverse relationship between log urine fluoride and IQ was observed; however, this significant relationship was not observed in children with the CC (wild-type) or CT (hybrid) genotypes (Cui *et al.* 2018).

Figure 7. Number of Lower Risk-of-bias Studies that Evaluated Thyroid Hormones in Children and Adults by Endpoint and Direction of Effect*

Endpoint	Direction of Effect ↑	Direction of Effect NS	Grand Total
serum T3	1	3	4
serum T4		4	4
serum TSH	5	2	7
self-reported thyroid condition		1	1
thyroid disease		1	1
hypothyroidism prevalence	1		1

*Interactive figure and additional study details in Tableau® (https://public.tableau.com/profile/ntp.visuals#!/vizhome/Fluoride_EpiThyroid_UPDATE/Figures6and7). This figure displays study counts for lower risk-of-bias studies in both children and adults, as these counts are most relevant to the summary of fluoride-related mechanistic effects in lower risk-of-bias studies. Counts for higher risk-of bias studies and studies by age (i.e., children, adults, or children/adults combined) can also be accessed in the interactive figure in Tableau®. Study counts are tabulated by significance (unless if study footnotes in Tableau indicate that statistical significance was not tested)—statistically significant increase (↑), statistically significant decrease (↓), or not significant (NS). For example, the "↑" column displays numbers of unique studies with significantly increased results.

Figure 8. Number of Higher Risk-of-bias Studies that Evaluated Thyroid Hormones in Children by Endpoint and Direction of Effect*

Endpoint	Direction of Effect ↑	Direction of Effect ↓	Direction of Effect NS	Grand Total
serum T3	1	1	6	8
serum T4	3		5	8
serum TSH	6	2	1	9

*Interactive figure and additional study details in Tableau® (https://public.tableau.com/profile/ntp.visuals#!/vizhome/Fluoride_EpiThyroid_UPDATE/Figures6and7). This figure displays study counts for higher risk-of-bias studies in children, as these counts are most relevant to the summary of fluoride-related effects on thyroid hormones in higher risk-of-bias studies. Counts for lower risk-of bias studies, studies in adults, or all studies combined, can also be accessed in the interactive figure in Tableau®. Study counts are tabulated by significance (unless if study footnotes in Tableau indicate that statistical significance was not tested)—statistically significant increase (↑), statistically significant decrease (↓), or not significant (NS). For example, the "↑" column displays numbers of unique studies with significantly increased results.

Animal Learning and Memory Data

[Note: An earlier version of the monograph underwent NASEM committee peer review in November 2019 (NASEM 2020). In this earlier review the committee criticized this section primarily over concerns that the NTP's risk-of-bias evaluations failed to adequately capture a number of important threats to internal validity that are specific to neurobehavioral outcomes in animal tests. The committee found

several examples of studies cited in the monograph where national or international guidelines for performance or statistical analyses were not followed or descriptions of methods were insufficient to evaluate their adequacy. Additionally, the committee took issue with a conclusion of the prior peer reviewed systematic review of the experimental animal literature (NTP 2016) concerning the degree to which motor activity deficits might compromise neurobehavioral assessments, affecting the directness of applicability of deficits in animal learning and memory to support the plausibility of IQ deficits in exposed children.

The NTP generally does not take issue with the NASEM peer review comments and acknowledges that further efforts to disentangle the potential for motor activity deficits to influence tests of learning and memory in the fluoride literature are warranted. The NTP agrees with the comments of the NASEM committee concerning the overall poor quality of the experimental animal database, with many studies suffering from major reporting deficiencies. NTP also found these to be general issues with the experimental animal database and were identified as deficiencies that led to the inadequate conclusion. However, the following experimental animal study section remains largely unchanged from the initial version of the monograph reviewed by NASEM in November 2019. The reasons for this are: (1) because a more critical risk-of-bias assessment would result in fewer relevant animal studies judged to be of high quality; (2) because the highest quality experimental animal study reviewed for this monograph (McPherson *et al.* 2018) did not find effects of fluoride on learning, memory or motor activity in the critical $\leq$20 ppm in drinking water concentration range; and (3) because of the availability of a large number of human epidemiology studies directly addressing neurobehavioral and cognitive effects of fluoride in children, a decision was made to focus efforts to address comments on the critical human epidemiology evaluation in this revised and updated monograph. NTP acknowledges the helpful comments of the NASEM committee on the following section and refers readers to the NASEM review (NASEM 2020) when considering the information provided. For the purpose of this updated review the NTP considers that the experimental animal data remain *inadequate* to inform conclusions on whether fluoride exposure is associated with cognitive effects (including cognitive neurodevelopmental effects) in humans. NTP is aware of a number of additional relevant experimental animal studies published since the literature cutoff date for the monograph. These additional studies have not been formally reviewed and may shed further light on these issues].

In 2016, NTP conducted a systematic review of the available experimental animal studies to develop level-of-evidence conclusions on the association between fluoride exposure and neurobehavioral effects, specifically effects related to learning and memory impairment (NTP 2016). As previously discussed, the evaluation of the animal body of evidence in this assessment is an extension of the NTP (2016) systematic review and is consistent with the methodology and format used in that report.

NTP (2016) identified two main issues with the animal body of evidence related to effects of fluoride exposure on learning and memory: indirectness and concerns for risk of bias. The concern related to indirectness was based on the fact that many learning and memory tests rely on a motor response (e.g., latency to achieve the desired effect). Changes in motor function or activity levels associated with fluoride exposures could complicate the interpretation of the results on learning and memory test performance depending on the outcome measured. The directness of the measure as an indicator of learning and memory (i.e., the ability to rule out impaired motor or sensory function) was considered when addressing confidence in the data. Concerns in these studies related to risk of bias included the following factors: lack of randomization, lack of blinding or other methods to reduce potential bias at outcome, lack of exposure information, lack of control for litter effects, lack of expected response in the

control animals, and lack of reporting of other key study information such as sample size or sex of the animals.

Since the NTP (2016) report was published, additional experimental animal studies were identified that evaluated learning and memory impairment associated with fluoride exposure, including 12 developmental exposure studies (Banala and Karnati 2015, Mesram *et al.* 2016, Zhu *et al.* 2017, Sun *et al.* 2018, Wang *et al.* 2018, Ge *et al.* 2018b, McPherson *et al.* 2018, Zhao *et al.* 2019, Ge *et al.* 2018a, Chen *et al.* 2018, Bartos *et al.* 2018, Banala *et al.* 2018); 5 Morris water maze study in adults (Zheng *et al.* 2016, Niu *et al.* 2018, Dong *et al.* 2017, Sharma *et al.* 2018, Yang *et al.* 2018); and 7 other maze studies in adults (Pulungan *et al.* 2016, Shalini and Sharma 2015, Sharma *et al.* 2018, Sudhakar *et al.* 2017, Nageshwar *et al.* 2018, Yuan *et al.* 2019, Raju *et al.* 2019). In addition, 12 studies were identified that evaluated motor activity/coordination or sensory effects without evaluating learning and memory impairment (Adedara *et al.* 2017a, Nageshwar *et al.* 2017, Nkpaa and Onyeso 2018, Sudhakar *et al.* 2018b, Ahmad *et al.* 2017, Kinawy and Al-Eidan 2018, Manusha *et al.* 2019, Sudhakar *et al.* 2018a, Agustina *et al.* 2018, Lu *et al.* 2019, Jia *et al.* 2019, Li *et al.* 2019).

Although Adedara *et al.* (2017a) and Nkpaa and Onyeso (2018) evaluated exploration, the authors concluded that the track plots in the open field novel environment test were consistent with impaired locomotor activity in the fluoride-treated animals. The additional studies reviewed did not address the concern of indirectness and most included risk-of-bias concerns; however, a few of these more recent studies are notable in that they provide results on learning and memory effects that could possibly be distinguished from effects on motor activity. Bartos *et al.* (2018) used a step-down inhibitory avoidance test to evaluate short-term and long-term memory in rat offspring. Although the authors did not discuss activity in the animals, this test would be expected to result in increased latency in animals if there was decreased activity with fluoride exposure. The fluoride-treated female offspring, however, had decreased latency indicating diminished memory of the foot shock. Chen *et al.* (2018) also evaluated female rat offspring (treatment continued until the offspring were 6 months old) and observed an effect of fluoride on latency to reach the platform and the number of platform crossings in the Morris Water Maze; however, swimming speed was measured, and no changes were observed. The tracks during the spatial probe test were also very different in the two higher exposed groups (i.e., 50 and 100 mg/L NaF), suggesting that the animals did not know the location of the platform. It is not clear if litter effects were addressed in the study.

After further evaluation of the data available in NTP (2016) and in this update, it is concluded that the animal data are <u>inadequate</u> to evaluate the effects of fluoride on learning and memory primarily due to the inability to separate the learning and memory effects from the effects on motor activity or motor coordination. The majority of the studies that evaluated effects of fluoride on learning and memory did not also evaluate a motor activity component to determine if the learning and memory effects could be attributed to motor activity or coordination deficits. Of the studies that did evaluate both learning and memory and motor activity/coordination, studies mainly found an association between fluoride exposure and both types of neurological outcomes or found no effect of fluoride exposure on either type of neurological outcome irrespective of the dose range or duration of dosing. In addition, studies that found effects on motor activity/coordination or learning and memory often did not provide sufficient indicators of general health of the animals to reliably attribute impaired performance on a task to a specific acquisition of learning and memory or motor activity/coordination. The few studies that provided this information used different test methods or results were inconsistent. Thus, it is difficult to conclude that evidence from experimental animal studies is meaningful when considering the specific question of fluoride's potential influence on human IQ or cognitive function, particularly at

human -relevant exposure levels. Based on this consideration, the experimental animal body of evidence does not contribute to confidence in conclusions derived from human epidemiological studies with respect to effects on human IQ. Although the evidence supports an association between fluoride exposure and neurodevelopmental effects, the data are not sufficient to support the primary effect evaluated in children (i.e., IQ) nor is it sufficient to support a conclusion on cognitive effects in adults especially in the absence of additional adult human data.

Mechanistic Data in Animals

There are a wide variety of studies in animals that evaluate mechanistic effects potentially related to neurological changes following oral fluoride exposure (see **Figure 9**). Categories of mechanistic endpoints with the largest amount of available data include changes in biochemical components of the brain or neurons, neurotransmitters, oxidative stress, histopathology, and thyroid function. Limiting the data to studies with at least one exposure at or below 20 ppm fluoride drinking water equivalents (gavage and dietary exposures were back calculated into equivalent drinking water concentrations for comparison) still provided a sufficient number of studies for evaluation of these mechanistic endpoints. Neurotransmitter and biochemical changes in the brain and neurons were considered to be the mechanistic areas with the greatest potential to demonstrate effects of fluoride on the brain of animals in the lower dose range and provide evidence of changes in the brain that may relate to lower IQ in children (see **Figure 10**). Histological data can be useful in determining whether effects are occurring in the brain at lower fluoride concentrations; however, author descriptions of these effects may be limited thereby making it difficult to directly link histological changes in the brain to learning and memory effects. Oxidative stress is considered a general mechanistic endpoint that cannot be specifically linked to neurodevelopmental or cognitive effects in humans; however, like histopathology, it may help in identifying changes in the brain occurring at lower concentrations of fluoride. Although any effects in the brain or neurological tissue at lower concentrations of fluoride may support reduced IQ in humans, it may be difficult to distinguish the potential effects of fluoride on learning and memory functions from other neurological or general health outcomes.

Figure 9. Number of Animal Mechanistic Studies for Fluoride by Mechanistic Category and Exposure Level*

Mechanism	Dose Level	
	All	≤20 ppm
Biochemical (brain/neurons)	66	25
Neurotransmitters	53	23
Oxidative stress	78	25
Histopathology	67	30
Apoptosis/cell death	19	7
Inflammation	7	5
Thyroid	50	17
Other	3	2

*Interactive figure and additional study details in Tableau® (https://public.tableau.com/profile/ntp.visuals#!/vizhome/Animal_Mechanisms_All_June2019/Figure8). The number of

studies that evaluated mechanistic effects associated with at least one exposure at or below 20 ppm fluoride is tabulated in the "≤20 ppm" column. The total number of studies per mechanistic category are summarized in the "All" column.

The following sections summarize the mechanistic data by category of mechanistic endpoint. Although there is some evidence of consistency in mechanistic effects, overall these data are insufficient to increase confidence or support a change to hazard conclusions.

Neurotransmitters

Twenty of 23 neurotransmitter studies assessed changes in brain cholinesterase activity associated with fluoride exposure at or below 20 ppm fluoride. Acetylcholine is a major neurotransmitter involved in learning, memory, and intelligence (Chen 2012, Gais and Schonauer 2017). AChE is responsible for the breakdown of acetylcholine in the synapses of nerve cells. Changes in cholinesterase, acetylcholine, or AChE could be related to effects on memory. Evidence of an effect varied among the lower risk-of-bias studies that assessed changes in cholinesterase or acetylcholine (n = 11 drinking water studies) (Gao *et al.* 2009, Baba *et al.* 2014, Adedara *et al.* 2017a, Khan *et al.* 2017, Gao *et al.* 2008a, Akinrinade *et al.* 2015a, Sun *et al.* 2000 [translated in Sun *et al.* 2008], Chouhan *et al.* 2010, Mesram *et al.* 2016, Liu *et al.* 2010, Nkpaa and Onyeso 2018), with the majority of studies reporting evidence of an effect that is considered inconsistent with the phenotypic outcome. Decreases in cholinesterase will cause increases in acetylcholine, which can have a positive effect on learning and memory; however, long-term decreases in cholinesterase can lead to secondary neuronal damage occurring in the cholinergic region of the brain (Chen 2012).

Five of the 11 studies with lower risk of bias (Gao *et al.* 2009, Baba *et al.* 2014, Adedara *et al.* 2017a, Khan *et al.* 2017, Nkpaa and Onyeso 2018) found statistically significant decreases in cholinesterase or AChE in brain homogenates (with some brains dissected into specific regions prior to homogenizing) with fluoride concentrations in drinking water at or below 20 ppm, and 4 of the 5 studies found statistically significant decreases in cholinesterase or AChE below 10 ppm. The 5 studies were conducted in rats (Wistar or Sprague-Dawley) with exposure ranging from 28 days to 6 months. An additional 2 out of 11 studies (Gao *et al.* 2008a, Akinrinade *et al.* 2015a) reported decreases in brain homogenate AChE at concentrations at or below 20 ppm fluoride in drinking water, but statistical significance was not reached. These studies were also conducted in rats with exposure for 30 days or 3 months. Gao *et al.* (2008a) reported a dose-dependent decrease in brain homogenate AChE in the low (5 ppm fluoride) and high (50 ppm fluoride) treatment groups compared with the control group, but the decrease was only statistically significant in the high dose group. Similarly, Akinrinade *et al.* (2015a) observed a dose-dependent decrease in percent intensity of AChE immunohistochemistry in the prefrontal cortex associated with 2.1 and 10 ppm sodium fluoride in the drinking water, but neither result was statistically significant. Gao *et al.* (2009) found lower brain homogenate AChE levels in the 5-ppm animals compared with the 50-ppm animals; therefore, the results were not always dose dependent.

Relative to the above-mentioned studies, 2 of the 11 lower risk-of-bias studies observed opposite effects on brain cholinesterase levels. Sun *et al.* (2000) [translated in Sun *et al.* 2008] observed a significant increase in brain cholinesterase in Kunming mice associated with fluoride drinking water concentrations from 10 to 100 mg/L, but did not observe a dose response. Chouhan *et al.* (2010) did observe a dose-related increase in AChE levels in brain homogenate of Wistar rats with sodium fluoride concentrations of 1 to 100 ppm for 12 weeks and noted statistically significant results at 1, 50, and 100 ppm but not at 10 ppm.

Mesram *et al.* (2016) did not assess changes in AChE but observed a significant decrease in acetylcholine levels in cerebral cortex homogenate through 30 days of age in rats treated in utero with 20 ppm

sodium fluoride, which may suggest an increase in AChE levels. Likewise, Liu *et al.* (2010) did not assess changes in AChE, but measured nicotinic acetylcholine receptors (nAChRs) in brain homogenate of rats following drinking water fluoride exposure, which the authors stated could modulate physiological and pharmacological functions that are involved in learning and memory-related behaviors. Significant decreases in the protein expressions of nAChR subunits at 2.26 ppm fluoride were observed; however, the corresponding receptor subunit mRNAs did not exhibit any changes (Liu *et al.* 2010).

The studies that assessed other neurotransmitters of the brain and neurons were too heterogeneous or limited in number to make any determination on mechanism, even before limiting the review of the data to lower risk-of-bias studies. There were only five studies that evaluated dopamine and/or metabolites (Tsunoda *et al.* 2005, Chouhan *et al.* 2010, Reddy *et al.* 2014, Banala *et al.* 2018, Sudhakar and Reddy 2018). Four of the studies observed decreases in dopamine levels in the brain with exposures less than 20 ppm fluoride (Reddy *et al.* 2014, Chouhan *et al.* 2010, Banala *et al.* 2018, Sudhakar and Reddy 2018); however, the fifth study (Tsunoda *et al.* 2005) observed increased dopamine and metabolites at fluoride exposures below 20 ppm (with statistical significance achieved only for the metabolite homovanillic acid in one brain region). No differences from the control group were observed at levels above 20 ppm fluoride. Other neurotransmitters were evaluated at or below 20 ppm fluoride exposure, but generally only in a couple of studies.

Biochemistry (brain/neurons)

Similar to above, the endpoints measured in brain biochemistry studies were too heterogeneous or limited in number to make any determination on potential relevance of mechanism, even before limiting the review of the data to lower risk-of-bias studies (see **Figure 10**). Endpoints related to biochemical changes in the brain or neurons included carbohydrate or lipid changes, RNA or DNA changes, changes in gene expression, or changes in protein expression. For the most part, only a single study was available for any given endpoint. The largest body of evidence on biochemistry was on protein level in various brain regions. Eleven lower risk-of-bias studies were identified that evaluated protein levels; however, few studies evaluated the same proteins or areas of the brain. In the few cases where the same protein was evaluated, results were not always consistent. These data are insufficient to increase confidence or support a change to hazard conclusions.

Histopathology

Histopathology of the brain was evaluated in 31 studies with concentrations at or below 20 ppm fluoride, of which 15 studies had a lower potential for bias (Adedara *et al.* 2017b, Akinrinade *et al.* 2015a, Bhatnagar *et al.* 2002, Bhatnagar *et al.* 2011, Chouhan *et al.* 2010, Guner *et al.* 2016, Jiang *et al.* 2014, Lou *et al.* 2013, McPherson *et al.* 2018, Mesram *et al.* 2016, Niu *et al.* 2018, Pulungan *et al.* 2016, Nageshwar *et al.* 2018, Zhao *et al.* 2019, Jia *et al.* 2019). In all but one lower risk-of-bias study [Pulungan *et al.* (2016); gavage], animals were exposed to fluoride via drinking water. All lower risk-of-bias studies were conducted in rodents, and all but three studies were conducted in rats (Wistar [seven studies]; Sprague-Dawley [four studies]; Long-Evans hooded [one study]). Overall, the lower risk-of-bias studies that evaluated histopathology in the brain had low potential for bias for key questions regarding randomization and exposure characterization; however, eight studies were rated as probably high risk of bias for the key risk-of-bias question regarding outcome assessment based on lack of reporting of blinding of outcome assessors and/or inadequate description of outcome measures or lesions. Moreover, low image quality in some of the studies hampered the ability to verify the quality of the data. Further technical review of the 15 lower risk-of-bias studies was conducted by a board-certified pathologist. Based on confidence in the results for each study, the technical reviewer further categorized the lower risk-of-bias studies as studies with higher or lower confidence in the outcome

assessment, which is reflected in the following summary of the brain histopathology results. Main limitations of the histopathology data identified by the pathologist included lack of information on methods of euthanasia and fixation. Perfusion fixation is generally considered the best practice for lesions of the central nervous system in addition to complete fixation of the brain prior to its removal from the skull (Garman *et al.* 2016). Four of the lower risk-of-bias studies reported that they used this method (Bhatnagar *et al.* 2002, Bhatnagar *et al.* 2011, McPherson *et al.* 2018, Pulungan *et al.* 2016). Two of the lower risk-of-bias studies handled the brains before fixation was complete, which can produce artifacts that can resemble dead neurons (Zhao *et al.* 2019, Nageshwar *et al.* 2018). Fixation and brain removal details were inadequately described in the remaining lower risk-of-bias studies.

Although there was heterogeneity in the endpoints reported (e.g., cell size, shape, and counts; nuclei fragmentation; increased vacuolar spaces) and some variation in the consistency of the evidence based on the area of the brain evaluated, the majority of the lower risk-of-bias studies (11 of 14 drinking water studies) found some histological change in the brain of rats or mice treated with fluoride at concentrations at or below 20 ppm, of which 8 studies reported histological changes in the brain at or below 10 ppm. Histological changes in the hippocampus (one of the areas of the brain most evaluated for histological changes) associated with fluoride exposures at or below 20 ppm were reported in three of four lower risk-of-bias studies with higher confidence in the outcome assessment (Bhatnagar *et al.* 2002, Bhatnagar *et al.* 2011, Guner *et al.* 2016) and in three of four lower risk-of-bias studies with lower confidence in the outcome assessment (Jiang *et al.* 2014, Niu *et al.* 2018, Nageshwar *et al.* 2018). McPherson *et al.* (2018) was the only drinking water study (with higher confidence in the histopathology outcome assessment) that did not observe any histological changes in hippocampus at 10 or 20 ppm fluoride in male Long-Evans hooded rats exposed in utero through adulthood (>PND80). Although there are too few studies to definitively explain the inconsistency in results, McPherson *et al.* (2018) also did not observe any associations between fluoride exposure and impairments to learning and memory, which is inconsistent with the majority of developmental exposure studies that observed learning and impairments associated with fluoride exposure for other strains of rats. Similarly, histological changes in the cortex were reported in three of the four lower risk-of-bias drinking water studies with higher confidence in the outcome assessment (Chouhan *et al.* 2010, Bhatnagar *et al.* 2011, Akinrinade *et al.* 2015a) and in three of four lower risk-of-bias studies with lower confidence in the outcome assessment (Lou *et al.* 2013, Mesram *et al.* 2016, Nageshwar *et al.* 2018).

Histological changes were also consistently reported in other areas of the brain in studies with higher confidence in the outcome assessment, including the amygdala, caudate putamen, cerebellum, and hypothalamus, although each of these areas of the brain were only evaluated in one lower risk-of-bias study (Bhatnagar *et al.* 2011, Guner *et al.* 2016). Pulungan *et al.* (2016), one of two lower risk-of-bias studies with higher confidence in the outcome assessment that did not report histological changes in the brain, observed a decreasing trend in the number of pyramidal cells in the prefrontal cortex with increasing dose, but this was not changed at concentrations below 20 ppm (study administered sodium fluoride via gavage; the 5-mg/kg-day dose was considered to be equivalent to 15.3 ppm fluoride in drinking water) nor were any of the results statistically significant.

Oxidative stress

Oxidative stress in the brain was evaluated in 25 studies that examined concentrations at or below 20 ppm fluoride, of which 15 studies had lower potential for bias (Adedara *et al.* 2017a, Adedara *et al.* 2017b, Akinrinade *et al.* 2015b, Chouhan *et al.* 2010, Gao *et al.* 2008b, Guner *et al.* 2016, Mesram *et al.* 2016, Nkpaa and Onyeso 2018, Shan *et al.* 2004, Zhang *et al.* 2015a, Chouhan and Flora 2008, Gao *et al.* 2009, Khan *et al.* 2017, Bartos *et al.* 2018, Nageshwar *et al.* 2018). All of the lower risk-of-bias studies

were conducted in rats (mainly Wistar or Sprague-Dawley) and administered fluoride via drinking water with exposure durations ranging from 28 days to 7 months. Although there was heterogeneity in the endpoints reported (i.e., varying measures of protein oxidation, antioxidant activity, lipid peroxidation, and reactive oxygen species [ROS]) and some variation in the consistency of the evidence based on the endpoint, the majority of the studies (13 of 15 studies) (Adedara *et al.* 2017a, Adedara *et al.* 2017b, Akinrinade *et al.* 2015b, Gao *et al.* 2008b, Gao *et al.* 2009, Guner *et al.* 2016, Mesram *et al.* 2016, Nkpaa and Onyeso 2018, Shan *et al.* 2004, Zhang *et al.* 2015a, Khan *et al.* 2017, Nageshwar *et al.* 2018, Bartos *et al.* 2018) found evidence of oxidative stress in the brains of rats treated with fluoride at concentrations at or below 20 ppm, of which 10 studies reported oxidative stress in the brain below 10 ppm fluoride. The most consistent evidence of oxidative stress in the brain was reported through changes in antioxidant activity. Eleven of the 12 lower risk-of-bias studies that evaluated antioxidant activity reported an effect at concentrations at or below 20 ppm (Adedara *et al.* 2017a, Adedara *et al.* 2017b, Akinrinade *et al.* 2015b, Gao *et al.* 2008b, Gao *et al.* 2009, Guner *et al.* 2016, Mesram *et al.* 2016, Nkpaa and Onyeso 2018, Khan *et al.* 2017, Bartos *et al.* 2018, Nageshwar *et al.* 2018). Decreases in antioxidant activity using measures of superoxide dismutase (SOD) activity were reported in seven of eight lower risk-of-bias studies (Adedara *et al.* 2017a, Adedara *et al.* 2017b, Akinrinade *et al.* 2015b, Mesram *et al.* 2016, Nkpaa and Onyeso 2018, Khan *et al.* 2017, Nageshwar *et al.* 2018) and, among these seven studies, all that also measured changes in catalase (CAT) activity (n = 6 studies) also reported decreased activity (Adedara *et al.* 2017a, Adedara *et al.* 2017b, Mesram *et al.* 2016, Nkpaa and Onyeso 2018, Khan *et al.* 2017, Nageshwar *et al.* 2018). A decrease in total antioxidant capacity (T-AOC) as a measure of antioxidant activity was also consistently reported in two lower risk-of-bias studies (Gao *et al.* 2008b, Gao *et al.* 2009), and a decrease in glutathione peroxidase (GPx) activity was reported in two of three lower risk-of-bias studies (Adedara *et al.* 2017b, Nkpaa and Onyeso 2018).

Relative to the above-mentioned studies, 2 of the 15 lower risk-of-bias studies (Chouhan and Flora 2008, Chouhan *et al.* 2010) did not observe statistically significant effects on oxidative stress in the brain with concentrations at or below 20 ppm fluoride; however, the measure of oxidative stress evaluated in Chouhan and Flora (2008) and Chouhan *et al.* (2010) (glutathione [GSH] to oxidized glutathione [GSSG] ratio as an indication of antioxidant activity and ROS levels) were not evaluated in any other lower risk-of-bias study. Chouhan and Flora (2008) observed a dose-dependent increase in ROS levels associated with 10, 50, and 100 mg/L sodium fluoride in the drinking water; however, results were not statistically significant at any dose. In Chouhan *et al.* (2010), the levels of ROS were significantly higher at 50 ppm sodium fluoride in drinking water, but statistical significance was not met at doses below 20 ppm fluoride (1 and 10 ppm sodium fluoride) or at 100 ppm sodium fluoride; yet, hydrogen peroxide levels as a measure of ROS were found to be significantly increased at 15 ppm sodium fluoride in drinking water in studies conducted by another group of authors (Adedara *et al.* 2017a, Adedara *et al.* 2017b).

Apoptosis/cell death

Seven lower risk-of-bias studies were identified that evaluated apoptosis with concentrations at or below 20 ppm fluoride. Results from these studies were inconsistent and were insufficient for evaluating fluoride-induced apoptosis. These data are insufficient to increase confidence or support a change to hazard conclusions.

Inflammation

Five lower risk-of-bias studies were identified that evaluated potential effects of fluoride on inflammation with concentrations at or below 20 ppm. The inflammation markers were too heterogeneous or limited in number to make any determination on potential relevance of mechanism,

even before limiting the review of the data to lower risk-of-bias studies. These data are insufficient to increase confidence or support a change to hazard conclusions.

Thyroid

Seventeen studies were identified that evaluated potential effects of fluoride on the thyroid with concentrations at or below 20 ppm (see **Figure 9**). These animal thyroid data are not further described because this endpoint has been directly evaluated in a number of human studies that have failed to identify consistent evidence to suggest that thyroid effects are a requisite mechanism by which fluoride causes neurodevelopmental or cognitive effects in humans.

Figure 10. Number of Lower Risk-of-bias Animal Studies that Evaluated Biochemical, Neurotransmission, and Oxidative Stress Effects at or Below 20 ppm by Mechanism Subcategory and Direction of Effect*

Mechanism	Mechanism Subcategory	Direction of Effect ↑	↓	NS	Grand Total
Biochemistry (brain/neurons)	Carbohydrate/lipid-related		1	1	2
	Gene expression		1		1
	Protein levels associated with brain function	8	7	7	11
	Other biochemical	2			2
Neurotransmitters	Cholinesterase	2	7	3	11
	Dopamine and metabolites		1		1
	Other neurotransmitters	2	2	1	2
Oxidative stress	Antioxidant activity	1	10	4	12
	Lipid peroxidation byproduct	7		4	11
	Protein oxidation	2		1	2
	ROS	2		2	4

*Interactive figure and additional study details in Tableau® (https://public.tableau.com/profile/ntp.visuals#!/vizhome/Fluoride_Animal_SelectMechanisms_UPDATE/Figure9). This figure displays study counts for lower risk-of-bias studies, as these counts are most relevant to the text in this section. Counts for higher risk-of bias studies or all studies combined can be accessed in the interactive figure in Tableau®. Study counts are tabulated by significance—statistically significant increase (↑), statistically significant decrease (↓), or not significant (NS). For example, the "↑" column displays numbers of unique studies with at least one endpoint in the mechanistic subcategory with significantly increasing results at fluoride exposure levels of ≤20 ppm. These columns are not mutually exclusive (i.e., a study may report on multiple endpoints with varying results within a single mechanistic subcategory and therefore may be reflected in the counts for the "↑", "↓", and NS columns, but would only be counted once in the Grand Total column). Endpoints, species, strain, sex, and exposure duration are available for each study in the interactive figure in Tableau®.

In Vitro/Mechanistic Data on Neurodevelopmental or Cognitive Effects

Although in vitro data were collected as part of the systematic review process, NTP determined that the information on neurological effects obtained from these studies is too general, and results cannot necessarily be attributed to effects on learning and memory or other cognitive functions at this time. The in vitro data may help support specific mechanisms identified from in vivo mechanistic data; however, as described above, no specific mechanism has been determined for fluoride effects on learning and memory or other neurodevelopmental or cognitive outcomes.

Evidence Synthesis for Neurodevelopmental or Cognitive Effects

There is consistent evidence that exposure to fluoride is associated with cognitive neurodevelopmental effects in children. There is moderate confidence in the human data in children from 5 well-conducted

prospective studies, supported by 14 cross-sectional studies where exposure was identified as likely occurring prior to outcome. The human body of evidence in adults is considered inadequate to evaluate whether fluoride exposure is associated with cognitive effects due to low confidence in the human data in adults, a limited number of studies, and a lack of evidence of an effect (i.e., there is not sufficient evidence of an effect, but the confidence in the data is not high enough to conclude that there is no effect). The animal data are inadequate to evaluate for learning and memory effects primarily due to the uncertainty in distinguishing effects on cognitive outcomes from secondary effects on the nervous system or general health including motor activity issues; however, these data do provide evidence of other neurodevelopmental effects. There is also evidence from mechanistic studies of adverse neurological effects of fluoride in humans and animals of unknown relationship to cognition.

The initial moderate confidence is based on 19 studies where exposure occurred prior to outcome and that evaluated individual-based outcomes and used a comparison group. Factors considered for upgrading or downgrading the confidence are as follows:

- **Risk of bias**: Only studies that were considered to have lower risk of bias were included in the moderate confidence rating; therefore, there is no downgrade for risk-of-bias concerns.

- **Unexplained inconsistencies:** The data are relatively consistent and there was no downgrade for this factor. In terms of IQ data, 17 studies observed significant effects associated with fluoride, and 2 studies found no significant association but neither of these studies adjusted for confounders. Consistency among neurodevelopmental effects other than IQ was also considered; however, the conclusions are based on the IQ data so these other neurodevelopmental effects would not impact a potential adjustment in confidence. Studies measuring neurodevelopmental effects other than IQ did not show consistent effects. It is not known whether fluoride exposure would be expected to be associated with neurodevelopmental outcomes in addition to IQ or other cognitive measures.

- **Indirectness:** IQ in humans is a direct measure of effect and therefore no adjustment in confidence is warranted.

- **Imprecision:** The meta-analysis indicates that there was no reason to downgrade due to imprecision.

- **Publication bias:** While the meta-analysis that estimated the pooled SMD among 46 included studies (both higher and lower risk-of-bias) indicated that there was potential for publication bias, a subgroup analysis indicated that there was no publication bias among the lower risk-of-bias studies (see **Figure A5-8**). Among the higher risk-of-bias studies, the trim-and-fill analysis estimated that, in the absence of publication bias, the negative direction of effect and statistical significance remained (**Figure A5-9**). For the meta-analysis that calculated a pooled effect estimate among the studies with individual-level measures, the funnel plot indicated publication bias; however, the trim-and-fill analysis estimated that once adjusted for publication bias, the negative direction of effect remained (**Appendix 5**, **Figure A5-16** and **Figure A5-18**). Therefore, no downgrade was applied for publication bias.

- **Large magnitude of effect:** While some individual studies indicate a large magnitude of effect, the overall pooled effect estimate from the meta-analysis of studies with individual-level

measures does not demonstrate a large magnitude of effect (**Appendix 5**). Therefore, the overall data would not support an upgrade due to a large magnitude of effect.

- **Dose-response:** Linear dose-response models provide the best fit to the data in studies examining individual-level measures of fluoride exposure and IQ. A meta-analysis of studies that compared mean IQ scores between groups of children with different levels of fluoride exposure showed a significantly lower mean SMD at higher concentrations of fluoride (>1.5 mg/L) from water (SMD: −0.14; 95% CI: −0.19, −0.08; n = 31 studies) and urine (SMD: −0.18; 95% CI: −0.31, −0.05; n = 22 studies) (**Appendix 5**); however, the dose-response relationship at fluoride concentrations below 1.5 mg/L fluoride in urine or drinking water is less certain. The overall dose-response could be used to upgrade the confidence in the body of evidence.

- **Residual confounding:** Xiang *et al.* (2003a), Xiang *et al.* (2011), and Wang *et al.* (2012) studied the same population where arsenic occurred in the area with low fluoride, but did not occur in the area with high fluoride. This would have biased the results toward the null, but there was a significantly lower IQ scores in the area with high fluoride. The remaining studies do not provide enough information to consider residual confounding as an impactful factor for the body of evidence. Therefore, the overall data would not support an upgrade due to residual confounding.

- **Consistency:** There is consistent evidence across study populations and study designs that fluoride is associated with lower IQ scores at higher concentrations of fluoride. There is uncertainty and less of a consistent pattern at concentrations below 1.5 mg/L. There is also a lack of consistency observed with and among other types of neurodevelopmental effects. The consistency in the overall results of the data set could increase the confidence.

Summary judgement on potential upgrades or downgrades in the confidence: Although the OHAT approach for evidence integration allows for the initial confidence in the body of evidence to be increased based on consistency or dose response, the NTP judgement is that the magnitude of effect and the overall strength and quality of the human literature base provides a moderate confidence in the body of evidence that fluoride causes cognitive neurodevelopmental effects in children.

The moderate confidence in the body of evidence in children translates to a moderate level of evidence that fluoride is associated with lower IQ and other cognitive neurodevelopmental effects in children.

The limited and weaker evidence of cognitive effects in adults is considered to provide an inadequate level of evidence that fluoride is associated with cognitive effects in adults. The animal body of evidence is also considered to provide an inadequate level of evidence for cognitive effects in adults.

Integration of these level-of-evidence conclusions supports an initial hazard conclusion of *presumed to be a cognitive neurodevelopmental hazard to humans* because of the extent and consistency of effect in the available data in children. Because most of the available studies evaluated intelligence in children, the primary focus in human data was on IQ and other cognitive neurodevelopmental effects, which is the primary basis for the hazard conclusion. A separate conclusion on other neurodevelopmental effects was not reached based on limited information in humans.

The moderate level of evidence in the human data in children supports a hazard conclusion of *presumed* instead of *suspected* due to the relatively large and consistent body of evidence, especially in relation to measures of IQ (17 of 19 lower risk-of-bias studies that assessed IQ reported an association between

higher fluoride and lower IQ scores) across multiple populations. A conclusion of *presumed* is supported by a statistically significant effect observed in the meta-analysis. Furthermore, the *presumed* hazard conclusion is supported by the low expectation that new studies would decrease the hazard conclusion.

Effects in children

- **Human body of evidence:** Moderate Confidence = Moderate Level of Evidence
- **Animal body of evidence:** Overall poor quality of studies and few studies that specifically assess effects on learning and memory after exposure during developmental periods separately from other neurological effects including motor activity = Inadequate Level of Evidence
- **Initial hazard conclusion (Moderate Human x Inadequate Animal)** = Presumed to be a Cognitive Neurodevelopmental Hazard to Humans
- **Final hazard conclusion (after consideration of biological plausibility)** = Presumed to be a Cognitive Neurodevelopmental Hazard to Humans

Effects in adults

- **Human body of evidence:** Low Confidence with no discernible effect = Inadequate Level of Evidence
- **Animal body of evidence:** Overall poor quality of studies and few studies that specifically assess effects on learning and memory after exposure in adulthood separately from other neurological effects including motor activity = Inadequate Level of Evidence
- **Initial hazard conclusion (Inadequate Human x Inadequate Animal)** = Not classifiable
- **Final hazard conclusion (after consideration of biological plausibility)** = Not classifiable

Table 7. Neurodevelopmental and Cognitive Function Evidence Profile for Fluoride

INITIAL CONFIDENCE for each body of evidence (# of studies)	Factors decreasing confidence "---" if no concern; "↓" if serious concern to downgrade confidence					Factors increasing confidence "---" if not present; "↑" if sufficient to upgrade confidence				FINAL CONFIDENCE RATING
	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Large Magnitude	Dose Response	Residual Confounding	Consistency Species/Model	
*Human IQ or cognitive function tests in children**										
Initial Moderate (5 prospective cohort studies[a]; 14 cross-sectional studies[b])	---	---	---	---	---	---	---	---	---	**Moderate**
Initial Low (7 cross-sectional studies)[c]	---	---	---	---	---	---	---	---	---	**Low**
*Human IQ or cognitive function tests in adults***										
Initial Low (2 cross-sectional studies)[d]	---	---	---	---	---	---	---	---	---	**Low**
Animal learning and memory or cognitive function										
Inadequate to assess effects in human										

References:

Human: Barberio *et al.* (2017b)[c]; Bashash *et al.* (2017)[a]; Bashash *et al.* (2018)[a]; Choi *et al.* (2015)[b]; Cui *et al.* (2018)[c]; Cui *et al.* (2020)[c]; Das and Mondal (2016); Ding *et al.* (2011)[b]; Green *et al.* (2019)[a]; Jacqmin *et al.* (1994)[d]; Li *et al.* (2004) [translated in Li *et al.* 2008a] [b]; Li *et al.* (2016)[d]; Riddell *et al.* (2019)[c]; Rocha-Amador *et al.* (2007)[b]; Rocha-Amador *et al.* (2009)[b]; Saxena *et al.* (2012)[b]; Seraj *et al.* (2012)[b]; Soto-Barreras *et al.* (2019)[b]; Sudhir *et al.* (2009)[b]; Till *et al.* (2020)[a]; Trivedi *et al.* (2012)[c]; Valdez Jimenez *et al.* (2017)[a]; Wang *et al.* (2012)[b]; Wang *et al.* (2020b)[b]; Wang *et al.* (2020a)[c]; Xiang *et al.* (2003a)[b]; Xiang *et al.* (2011)[b]; Yu *et al.* (2018)[b]; Zhang *et al.* (2015b)[c]

*This includes learning disabilities, neonatal behavioral neurological assessment, mental development index, memory score for copy, and immediate recall.
**This includes Mini-Mental State Examination scores, psychomotor performance, and memory.

DISCUSSION

The overall objective of this evaluation was to undertake a systematic review of published literature to reach conclusions concerning the potential for exposure to fluoride to affect neurodevelopment and cognition. This review only addresses whether exposure to fluoride could present a potential hazard (i.e., has the potential to cause harm at any exposure level, including exposures that are higher than typically encountered from consuming fluoridated drinking water in the United States). Benefits of fluoride with respect to oral health are not addressed in this monograph.

Given this context, when focusing on human epidemiology studies with exposures in ranges typically found in the water distribution systems in the United States (0.7 mg/L for optimally fluoridated community water systems)[8] that can be evaluated for dose response, effects on cognitive neurodevelopment are inconsistent and therefore unclear. However, given the totality of the data, including studies with exposures to fluoride levels higher than the WHO safe water guideline of 1.5 mg/L in water (WHO 2011), the NTP concludes that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans.* This conclusion is based on a moderate level of evidence that shows a consistent and robust pattern of findings in human studies across several different populations demonstrating that higher fluoride exposure (e.g., >1.5 mg/L in drinking water) is associated with lower IQ and other cognitive effects in children. Limited and weaker evidence is considered to provide an inadequate level of evidence that fluoride is associated with cognitive effects in adults. The primary focus of the human data was on IQ and cognitive neurodevelopmental effects; therefore, the conclusion was based on these data. After further evaluation of the experimental animal data available in NTP (2016) and in this systematic review, NTP concludes that in terms of evaluating the effects of fluoride on learning and memory to support the cognitive effects observed in humans, the animal data are inadequate. The animal data do provide evidence for effects of fluoride on neurodevelopment; however, other neurodevelopmental outcomes were not further evaluated because of the limited information in humans. Biological plausibility of effects from mechanistic studies was considered but did not significantly influence the conclusions. Although multiple categories of mechanistic data were evaluated and provide some evidence of adverse effects in the brain, a coherent series of mechanistic events to account for fluoride-associated cognitive neurodevelopmental deficits is not sufficiently understood for these findings to contribute to the overall confidence assessment.

The human body of evidence provides a consistent and robust pattern of findings that higher fluoride exposure is associated with lower IQ scores in children. The moderate level of evidence is based on 5 lower risk-of-bias prospective cohort studies and 14 lower risk-of-bias cross-sectional studies that are considered to have sufficient evidence of fluoride exposure occurring prior to the outcome. The evaluation of the animal body of evidence in this assessment is an extension of the NTP (2016) systematic review on the association between fluoride exposure and neurobehavioral effects related to learning and memory in animals, which identified a concern related to indirectness. This concern was that many of the learning and memory tests rely on a motor response (e.g., latency to achieve the desired effect). The review of animal data published since the 2016 review focused on addressing this

[8]As of April 2020, 1.08% of persons living in the United States (~ 3.5 million people) were served by CWS containing ≥ 1.1 mg/L naturally occurring fluoride. CWS supplying water with ≥ 1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population (~ 1.9 million people), and systems supplying water with ≥ 2 mg/L naturally occurring fluoride served 0.31% of the U.S. population (~1 million people) (https://www.cdc.gov/fluoridation/data-tools/reporting-system.html).

indirectness concern. Further examination of the literature has not provided clarification of this issue. Due to the inability to separate these effects from effects on general health and other effects on the nervous system, the animal body of evidence is now considered inadequate to contribute to the evaluation of cognitive effects in humans. Although the animal data are not considered sufficient to specifically support the IQ changes observed in children, the data do support possible neurodevelopmental effects.

The NTP conclusion that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans* is supported by the extent, consistency, and robustness of the effect in the available data in children. Seventeen of the 19 lower risk-of-bias studies reported an association between higher fluoride exposure and lower IQ scores in children across multiple populations. Meta-analyses conducted at the recommendation of NASEM based on their review of the September 6, 2019 draft monograph provide further support for the hazard conclusion of *presumed* (NASEM 2020). The random-effects pooled SMD estimate from the 46 studies included in the group-level meta-analysis was consistent with two previous meta-analyses reporting statistically significant associations between higher fluoride exposure and lower IQ in children. A risk-of-bias subgroup analysis demonstrated that the significant negative association remained when the meta-analysis was restricted to 9 lower risk-of-bias studies. Further subgroup analyses by gender, age group, country, outcome assessment type, and exposure assessment type support a consistent and robust pattern of results. A second meta-analysis of the individual-level data from six lower risk-of-bias studies also provided evidence of a statistically significant negative association between fluoride exposure and lower IQ in children (overall pooled effect estimate per 1-mg/L increase in urinary fluoride was associated with a 1.40-point lower IQ score [95% CI: −2.33, −0.47]). Given the evidence, there is a low expectation that new studies would change the hazard conclusion.

There are few studies in humans and numerous studies in animals that evaluate mechanistic effects related to fluoride exposure. There are sufficient mechanistic data to determine that fluoride exposure at lower concentrations has effects on the nervous system; however, for the cognitive neurodevelopmental outcome evaluated, there are insufficient data to support a specific mechanism or mode of action. Due to the large number of mechanistic studies conducted in animals, evaluation of the mechanistic data in animals focused on studies that had exposures more relevant to humans (i.e., ≤20 ppm in the drinking water). Changes in AChE, which could potentially be related to cognitive effects such as IQ, were evaluated in one study of children and several studies of animals (measured in both the blood and in areas of the brain); however, the majority of these studies, including the study of children, reported results inconsistent with the phenotypic outcome. Animal studies that evaluated changes in other neurotransmitters and other biochemical measures provide some evidence of effects in the brain, but the data are limited due to the heterogeneity of the outcomes measured. Most consistently, studies evaluating histopathology and oxidative stress demonstrated that effects can occur in the brains of animals at or below 20 ppm, which, without supporting a specific mechanism or mode of action relevant to learning and memory impairments, provides evidence of an association between exposure to lower concentrations of fluoride and neurological effects in animals. Therefore, the evidence of neurological effects at exposure levels more relevant to humans that is demonstrated in the mechanistic data supports the NTP conclusion that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans;* however, it does not provide enough evidence to increase confidence in the human body of evidence or support a higher hazard identification conclusion.

Generalizability to the U.S. Population

For many years, fluoride concentrations were adjusted to levels between 0.8 and 1.2 mg/L in fluoridated community water systems in the United States. The U.S. Public Health Service recommended an adjustment downward to a fluoride concentration of 0.7 mg/L because of evidence of an increase in dental fluorosis in children (US DHHS 2015). In April 2020, the CDC Water Fluoridation Reporting System estimated that the majority (i.e., 97.5%) of fluoride concentrations in water for U.S. children and adolescents (≤19 years old) are below 1.2 mg/L (https://www.cdc.gov/fluoridation/data-tools/reporting-system.html).

NTP's conclusion that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans* is based on consistent evidence from 26 lower risk-of-bias studies that evaluated fluoride exposure and effects on children's IQ and other cognitive effects. Although there are many studies that evaluated associations between fluoride in the drinking water and IQ in children, no studies evaluating IQ were conducted in the United States. Generalizing the results from the IQ studies in this evaluation to the U.S. population can be difficult, in part because many studies were conducted in areas with fluoride drinking water concentrations that are much higher than drinking water fluoride concentrations in the United States. Among the human body of evidence evaluated for this assessment (including lower and higher risk-of-bias studies), there are 33 studies that evaluated associations between fluoride in drinking water and IQ in children and compared a reference or low exposure group to higher fluoride-exposed groups. Of these 33 studies, only 10 studies included fluoride exposure groups with fluoride concentrations < 1.5 mg/L (i.e., fluoride exposure groups that would potentially be relevant to levels observed in the United States) (Xu *et al.* 1994, Xiang *et al.* 2003a, Qin *et al.* 1990 [translated in Qin *et al.* 2008], Kang *et al.* 2011, Broadbent *et al.* 2015, Sebastian and Sunitha 2015, Sudhir *et al.* 2009, Zhang *et al.* 2015b, Zhang *et al.* 1998, Wang *et al.* 2020b). Of these 10 studies, 4 were considered to have lower risk of bias (Zhang *et al.* 2015b, Xiang *et al.* 2003a, Wang *et al.* 2020b, Sudhir *et al.* 2009).

In addition to the four studies mentioned above, several other studies that evaluated fluoride exposure on a continuous basis could be used to assess generalizability to the United States. This includes studies that examined fluoride exposure levels below 1.5 mg/L for which a dose response could be assessed. **Table 8** provides a summary of children's IQ studies that evaluated lower fluoride exposures (<1.5 mg/L) in drinking water and/or urine (assuming, for comparison purposes, an approximate 1-to-1 equivalence between drinking water fluoride and urinary fluoride concentrations) and provided information to evaluate dose response in the lower fluoride exposure range (e.g., three or more fluoride exposure groups or dose-response curve provided). Based on review of these studies (discussed further below), there is uncertainty if IQ changes in children occur at lower fluoride levels.

Among studies with lower risk of bias for which a dose response could be assessed, four of nine studies that examined fluoride exposure levels below 1.5 mg/L (**Table 8)** (Green *et al.* 2019, Zhang *et al.* 2015b, Xiang *et al.* 2003a, Wang *et al.* 2020b) applied regression models to individual exposure outcome measures and observed a linear association between urinary fluoride levels and lower IQ in children even at the lower fluoride concentrations. However, two of these studies (Xiang *et al.* 2003a, Zhang *et al.* 2015b) did not find an association between IQ and drinking water levels below 1.5 mg/L. Xiang *et al.* (2003a) observed a significantly lower IQ in endemic versus nonendemic villages, but when they grouped children from the endemic villages by exposure level, they did not observe a significantly lower IQ score for children exposed to lower mean exposure levels of fluoride (0.75 mg/L). Although a significant difference in IQ might not be expected due to the fact that there were only nine children in this group, the difference was less than one point in IQ. Zhang *et al.* (2015b) used a simple correlation

and did not observe a significant relationship between fluoride levels in the drinking water (with concentrations up to 1.57 mg/L) and IQ. Sudhir *et al.* (2009) observed a significant increase in IQ grade (which is associated with lower IQ) at concentrations of 0.7–1.2 mg/L. The other four of nine studies do not provide a clear dose response at the lower fluoride levels. Bashash *et al.* (2017) concluded that there was no clear association between IQ scores and maternal urinary fluoride below 0.8 mg/L. Yu *et al.* (2018) observed a correlation between lower IQ in children and fluoride exposure only with concentrations in drinking water above 3.4 mg/L or with urinary fluoride concentrations of 1.6 mg/L or higher. The study authors did note a decreased probability of having an IQ above 130 (i.e., 40% fewer people with high IQ for every 0.5-mg/L increase in fluoride) with water fluoride levels between 0.20 and 1.40 mg/L, but this was not observed with higher levels of fluoride. Although Cui *et al.* (2018) noted that IQ decreased in a "roughly linear manner" with increasing urinary fluoride, this is only apparent in the results for the TT genotype; based on the dose response, the authors concluded that the "safety threshold" was 1.73 mg/L. Ding *et al.* (2011) looked at mean differences for 10 different exposure groups and found notable decreases from the mean above approximately 1 mg/L.

Although there is less confidence in the findings from higher risk-of-bias studies, six studies identified with potential dose-response information demonstrated a similar uncertainty at the lower fluoride concentrations. Surprisingly, three of the studies (Aravind *et al.* 2016, Qin *et al.* 1990 [translated in Qin *et al.* 2008], Xu *et al.* 1994) found that the lowest IQ scores were in areas with the lowest and the highest fluoride concentrations. In these studies, the lowest fluoride concentrations ranged from 0.1–0.2 mg/L fluoride in Qin *et al.* (1990) [translated in Qin *et al.* 2008] to <1.2 mg/L in Aravind *et al.* (2016). Li *et al.* (1995) and Sebastian and Sunitha (2015) only observed lower IQ scores at concentrations above 2 mg/L. Das and Mondal (2016) found a steady decline in IQ with increasing urinary fluoride levels or exposure dose.

To further examine the dose response for lower IQ in the lower exposure region (e.g., <1.5 mg/L fluoride in drinking water or urine), a meta-analysis (**Appendix 5**) using a linear mixed model to analyze mean-effect estimates was performed. Twelve observations from 9 studies were included that reported one or more IQ measurements associated with drinking water fluoride exposures of < 1.5 mg/L and a reference group. This analysis did not show a statistically significant association with the mean SMD in children's IQ scores between exposed and reference groups (SMD = 0.32; 95% CI: −0.57, 1.20). A dose-response meta-analysis including seven observations from four studies with at least one exposure group < 1.5 mg/L urinary fluoride showed a non-statistically significant decrease in mean SMD (SMD = −0.13; 95% CI: −0.29, 0.03). Based on these results, effects of fluoride exposure on children's IQ at levels < 1.5 mg/L remain unclear and more studies at lower exposure levels are needed. A dose-response meta-analysis of studies with individual-level data could not be conducted due to the small number of studies (n = 10), the various types of exposure metrics, and the different types of reported effect estimates. More studies with lower levels of fluoride exposure from drinking water are still needed to fully understand potential effects at exposures in ranges typically found in the United States (i.e., <1.5 mg/L). Of note, the negative association between IQ and fluoride exposure via drinking water was statistically significant when extending the dose-response meta-analysis to include IQ measures from groups exposed to <2 mg/L in drinking water (SMD = −0.27; 95% CI: −0.36, −0.17). A statistically significant decrease in mean SMD in children's IQ scores was not seen in urinary fluoride measures of < 2 mg/L (SMD = −0.09; 95% CI: −0.22, 0.03) (**Table A5-3**).

When generalizing findings from the cited studies to exposures in the United States from fluoride in drinking water, it is important to consider that drinking water only comprises a portion of total exposures to fluoride. Although it can be assumed that children in all the studies cited in this document

are also exposed to fluoride from sources other than drinking water, these other exposures are likely to vary considerably depending on individual circumstances. Fluoride concentrations in drinking water alone do not reflect the magnitude of fluoride exposures to children who consume excessive amounts of fluoridated toothpaste or to formula-fed babies who consume powdered formula that is reconstituted with fluoridated water. A few studies also support the possibility of heightened sensitivities to the detrimental cognitive effects of fluoride exposures in individuals with certain polymorphisms in dopamine receptor D2, or catechol-O-methyltransferase (Cui *et al.* 2018, Zhang *et al.* 2015b), impacting dopamine catabolism and receptor sensitivity. Differential exposures to fluoride and genetic susceptibilities of children to fluoride appear to warrant further research.

Table 8. Human Studies with Lower Fluoride Exposure and Effects on IQ

Study	Exposure measures [mean ± SD (range)]	Notes
Lower risk-of-bias studies		
Bashash *et al.* (2017)	Maternal urine during pregnancy (mg/L) 0.90 ± 0.35 (0.23–2.36)	Authors concluded that the model suggested a nonlinear relationship with no clear association between IQ scores and maternal urine below 0.8 mg/L.
	Children's urine (mg/L) 0.82 ± 0.38 (0.18–2.8)	
Green *et al.* (2019)	Maternal urine during pregnancy (mg/L) 0.51 ± 0.36 (0.06–2.44) 0.40 ± 0.27 non-fluoridated areas 0.69 ± 0.42 fluoridated areas	Statistical methods indicated that including quadratic or natural-log effects of maternal urine or intake did not significantly improve the model. In addition, the authors tested separate models with two linear splines to see if the effect of maternal urinary fluoride or maternal fluoride intake significantly differed between lower and higher levels based on knots set at 0.5, 0.8, and 1.0 mg/L for urine and 0.4, 0.8, and 1 mg for intake. There were no differences.
	Maternal intake during pregnancy (mg/day) 0.54 ± 0.44 (0.01–2.65) 0.30 ± 0.26 non-fluoridated areas 0.93 ± 0.43 fluoridated areas	
	Drinking water (mg/L)* 0.31 ± 0.23 (0.04–0.87[1]) 0.13 ± 0.06 non-fluoridated areas 0.59 ± 0.08 fluoridated areas	
Cui *et al.* (2018)	Drinking water (mg/L)* 0.20–1.00 non-endemic 1.52–2.49 endemic	Study authors noted that the IQ decreased in a "roughly linear manner as the log-urine fluoride increased." TT genotypes of the dopamine receptor D2 gene had the strongest negative correlation between log-urine fluoride and IQ scores. The study authors determined a safety threshold of urine fluoride levels in subgroup TT as 1.73 mg/L. Drinking water fluoride levels were used to select children from different areas but were not used in the analysis.
	Children's urine Levels not provided; log-transformed with range of approximately −1.2–2.2	
Ding *et al.* (2011)	Drinking water (mg/L) * 1.31 ± 1.05 (0.24–2.84)	Although there was a significant correlation between urinary fluoride and IQ score, the main drop in IQ occurred at urinary fluoride levels of approximately 0.7–1.2 mg/L. At levels below 0.7 mg/L, data suggest a plateau with no apparent change in IQ compared with the mean. Drinking water fluoride levels were not used in the analysis.
	Children's urine (mg/L) 0.10–3.55	

Table 8. Human Studies with Lower Fluoride Exposure and Effects on IQ

Study	Exposure measures [mean ± SD (range)]	Notes
Sudhir *et al.* (2009)	Drinking water (mg/L)* <0.7 (level 1 villages) 0.7–1.2 (level 2 villages) 1.3–4.0 (level 3 villages) >4.0 (level 4 villages)	The number of intellectually impaired children gradually increased with increasing fluoride concentration in drinking water with an increase in IQ grade (which indicates a decrease in IQ) observed in the 0.7–1.2-mg/L villages. Children were placed in exposure groups based on Water Works Department records. Although children brought in water for verification of strata, it was not collected from all children but only from the first child using a different source of water. Therefore, these are considered group-level data. Note that all groups had a large proportion in the "intellectually impaired" category.
Wang *et al.* (2020b)	Drinking water (mg/L)* 1.39 ± 1.01 (0.20–3.90) Children's urine (mg/L) 1.28 ± 1.30 (0.01–5.54)	There was a significantly lower IQ in quartile 3 (1.00–1.90 mg/L) and quartile 4 (>1.90 mg/L), but not quartile 2 (0.70–1.00 mg/L) of drinking water. Urinary fluoride was only associated with a significantly lower IQ in quartile 4 (>2.28 mg/L).
Xiang *et al.* (2003a)	Drinking water (mg/L)* 0.36 ± 0.15 (0.18–0.76) non-endemic village 2.47 ± 0.79 (0.57–4.50) endemic village Endemic subgroups: group A: 0.75 ± 0.14 group B: 1.53 ± 0.27 group C: 2.46 ± 0.30 group D: 3.28 ± 0.25 group E: 4.16 ± 0.22 Children's urine (mg/L) 1.11 ± 0.39 (0.37–2.50) non-endemic village 3.47 ± 1.95 (0.90–12.50) endemic village	IQ in group A in the endemic village was not significantly lower than the non-endemic village, but IQ in all other groups was significantly lower. Although there were only 9 children in group A, the IQ difference was <1 point. Based on simple regression, there was a steady decline in IQ with increasing urinary fluoride.
Yu *et al.* (2018)	Drinking water (mg/L)* 0.50 ± 0.27 normal 2.00 ± 0.75 high Children's urine (mg/L) 0.41 ± 0.49 normal 1.37 ± 1.08 high	Study authors reported that participants' intelligence presented inverse non-linear dose-response relationships with fluoride content, with obvious decreases at relatively high level of fluoride exposure (drinking water fluoride levels at 3.4–3.90 mg/L and urinary fluoride levels at 1.60–2.50 mg/L). Study authors also note a decreased odds for having IQ ≥ 130 with drinking water fluoride at 0.20–1.40 mg/L (40% decrease with each 0.5-mg/L increase in fluoride), but not at higher concentrations.

Table 8. Human Studies with Lower Fluoride Exposure and Effects on IQ

Study	Exposure measures [mean ± SD (range)]	Notes
Zhang *et al.* (2015b)	Drinking water (mg/L)* 0.63 (0.58–0.68) control 1.40 (1.23–1.57) endemic fluorosis	There was a steady decline in IQ with increasing serum or urinary fluoride levels. A simple correlation did not find drinking water fluoride significantly correlated with IQ.
	Children's urine (mg/L) 1.1 ± 0.67 control 2.4 ± 1.01 endemic fluorosis	
	Children's serum (mg/L) 0.06 ± 0.03 control 0.18 ± 0.11 endemic fluorosis	
Higher risk-of-bias studies		
Aravind *et al.* (2016)	Drinking water (mg/L)* <1.2 low fluoride area 1.2–2 medium fluoride area >2 high fluoride area	Mean IQ scores (transformed into percentiles) were highest in the medium fluoride area for both boys and girls.
Das and Mondal (2016)	Groundwater (mg/L)* 2.11 ± 1.64 (0.25–9.40)	Based on simple regression, there was a steady decline in IQ with increasing urinary fluoride and increasing exposure dose. Groundwater fluoride levels were not used in the analysis but were used in calculating the children's exposure dose.
	Children's urine (mg/L) 0.45–17.00	
	Children's exposure dose (mg/kg-day) 0.017–0.203	
Li *et al.* (1995)	Children's urine (mg/L) 1.02 non-fluorosis area 1.81 slight fluorosis area 2.01 medium fluorosis area 2.69 severe fluorosis area	A significantly lower IQ score was observed in the medium and severe fluorosis areas compared to the non-fluorosis area. Children's urine was used as an individual measure of exposure to verify that the areas had different fluoride exposure levels; however, analysis was conducted based on residential area.
Qin *et al.* (1990) [translated in Qin *et al.* 2008]	Drinking water (mg/L)* 0.1–0.2 low fluoride area 0.5–1.0 normal fluoride area 2.1–4.0 high fluoride area	Average IQ scores (transformed into percentages) were significantly lower in both the low and high fluoride areas compared with the normal fluoride area.
Sebastian and Sunitha (2015)	Drinking water (mg/L)* 0.40 (low fluoride village) 1.2 (normal fluoride village) 2.0 (high fluoride village)	A significantly lower mean IQ score of children living in the high fluoride area compared with the low and normal fluoride areas was reported. Binary regression models using the low fluoride village as a reference observed an increased odds ratio (1.74) for increased IQ scores in the normal fluoride village and a decreased odds ratio (0.59) in the high fluoride village.
Xu *et al.* (1994)	Drinking water (mg/L)* 0.8 (control area) 0.38 (low fluoride area) 1.8 (high fluoride area)	Both low and high fluoride areas had IQ levels approximately 3 points below the IQ levels in the control area. There was no difference in IQ between the low and high fluoride areas.

*Data are group-level exposure data; exposure data without the asterisks are individual-level exposure data.

[1]Range data were obtained from Till *et al.* (2018).

The conclusion that fluoride is *presumed to be a cognitive neurodevelopmental hazard in children* is based on the consistency of the data; however, most lower risk-of-bias studies observed effects with drinking water concentrations above 1.5 mg/L. As noted above, describing the effects at 1.5 mg/L or below, which is more relevant to the exposures observed in the U.S. population, including from community water fluoridation, is more difficult. In the reviewed studies, when limiting studies to those that evaluated IQ at fluoride levels across a continuum that included the low dose range, results are less consistent.

Limitations of the Evidence Base

Few limitations exist in the lower risk-of-bias epidemiological studies used for the basis of the hazard conclusion. The main limitations in lower risk-of-bias epidemiological studies include:

- Few studies were available that assessed the association between fluoride exposure and the following:
    - Neurodevelopmental or cognitive effects in subjects from communities served by optimally fluoridated versus non-fluoridated water systems.
    - Neurobehavioral (i.e., cognitive) effects (particularly IQ) in adults.
    - Attention-related disorders including ADHD.
- Studies rarely separated the results by gender or provided information to indicate that gender was not a modifying factor, which limits the ability to evaluate how the association between fluoride exposure and cognitive neurodevelopmental effects in children might differ by gender.

Limitations in the epidemiological studies with higher risk of bias include:

- Many of the original publications were in a foreign language and provided limited details on methodology.
- Some studies lacked information regarding exposure and/or had serious limitations in the exposure assessment. Exposure assessment concerns include limited individual exposure information, a lack of information on fluoride sampling methods and timing of the exposure measurements, a lack of quantitation of levels of fluoride in drinking water, and a lack of individual-level information on fluorosis in areas reported to be endemic for fluorosis.
- The comparison groups in studies conducted in areas endemic for fluorosis may have still been exposed to high levels of fluoride or levels similar to those used in water fluoridation in the United States. This factor may have limited the ability to detect true effects.
- Most studies did not provide sufficient direct information (e.g., participation rates) to evaluate selection bias.
- Failure to address potential confounders was a main issue. Many studies conducted simple statistical analyses without accounting for any potential confounders. In cases where adjustments in analyses were made, often these studies did not account for potential confounders considered critical for that study population and outcome.

- Studies conducted in areas with high, naturally-occurring fluoride levels in drinking water often did not account for potential exposures to arsenic or iodine deficiencies in study subjects.
- Many studies did not account for potential exposures to lead as a residual confounder.
- Many studies lacked information on whether the outcome assessors were blind to the exposure group, including studies that examined children in their schools and subjects from high-fluoride communities.

Limitations in the animal studies include:

- The main limitation in the animal studies was the inability to separate possible learning and memory effects from effects on motor activity/coordination or sensory effects.
- Few learning and memory studies in animals evaluated motor activity or sensory effects. Studies that did evaluate motor activity or sensory effects often lacked discussion on general health of the animals when the endpoints measured could be affected by deficits in motor activity or sensory, such as latency to achieve a desired result.

A key data gap in the human and animal bodies of evidence includes the need for mechanistic insight into fluoride-related neurodevelopmental or cognitive changes.

Limitations of the Systematic Review

There are no major limitations of the systematic review. The human body of evidence included a large database of observational studies. Most of the observational studies were cross-sectional; however, nine of these were considered to be functionally prospective in nature. In addition, the systematic review covered a wide range of study designs, populations, and measures of fluoride exposure. The systematic review was designed to cover reports on all potential mechanistic data including effects on the thyroid. After review of the studies evaluating thyroid effects, studies that only evaluated goiters and other effects on thyroid size were not considered in this review. This is not considered a limitation because changes in thyroid size are not functional changes to the thyroid that could specifically indicate a mechanism for thyroid involvement in neurodevelopment. In addition, review of the mechanistic data was limited to in vivo studies with at least one concentration below 20 ppm. This is not considered a limitation for the systematic review since the mechanistic body of evidence was used to evaluate biological plausibility for the effects observed in humans; therefore, data were limited to concentrations that would be more reflective of human exposures. The decision to not more closely evaluate the in vitro data is not considered a limitation because there were sufficient in vivo data, and no key events were identified where in vitro data would provide additional insight.

CONCLUSION

Because the majority of available studies evaluated cognitive neurodevelopmental effects in children, the focus of the hazard conclusions is on cognitive neurodevelopmental effects, primarily IQ. When focusing on findings from studies with exposures in ranges typically found in drinking water in the

United States (0.7 mg/L for optimally fluoridated community water systems)[9] that can be evaluated for dose response, effects on cognitive neurodevelopment are inconsistent and, therefore, unclear. However, when considering all the evidence, including studies with exposures to fluoride levels higher than 1.5 mg/L in water, NTP concludes that fluoride is *presumed to be a cognitive neurodevelopmental hazard to humans*. This conclusion is based on a moderate level of evidence that shows a consistent and robust pattern of findings in human studies across several different populations demonstrating that higher fluoride exposure (e.g., >1.5 mg/L in drinking water) is associated with lower IQ and other cognitive effects in children. Limited and weaker evidence is considered to provide an inadequate level of evidence that fluoride is associated with cognitive effects in adults. The evidence from animal studies is inadequate to inform conclusions on cognitive effects, and the mechanisms underlying fluoride-associated cognitive neurodevelopmental effects are not well characterized.

[9]As of April 2020, 1.08% of persons living in the United States (~ 3.5 million people) were served by CWS supplying ≥ 1.1 mg/L naturally occurring fluoride. CWS supplying water at ≥ 1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population(~ 1.9 million people), and systems supplying water at ≥ 2 mg/L naturally occurring fluoride served 0.31% of the U.S. population (~1 million people) (https://www.cdc.gov/fluoridation/data-tools/reporting-system.html).

REFERENCES

Adedara IA, Abolaji AO, Idris UF, Olabiyi BF, Onibiyo EM, Ojuade TD, Farombi EO. 2017a. Neuroprotective influence of taurine on fluoride-induced biochemical and behavioral deficits in rats. *Chem Biol Interact* 261: 1-10.

Adedara IA, Olabiyi BF, Ojuade TD, Idris UF, Onibiyo EM, Farombi EO. 2017b. Taurine reverses sodium fluoride-mediated increase in inflammation, caspase-3 activity, and oxidative damage along the brain-pituitary-gonadal axis in male rats. *Can J Phys Pharmacol* 95: 1019-1029.

Agustina F, Sofro ZM, Partadiredja G. 2018. Subchronic administration of high-dose sodium fluoride causes deficits in cerebellar purkinje cells but not motor coordination of rats. *Biol Trace Elem Res* 188(2): 424-433.

Ahmad KR, Noor S, Jabeen S, Nauroze T, Kanwal MA, Raees K, Abbas T. 2017. Amelioration by jambul fruit extract of fluoride-induced hepato-nephronal histopathologies and impaired neuromotor capacity in mice. *Fluoride* 50: 2-14.

Akinrinade ID, Memudu AE, Ogundele OM. 2015a. Fluoride and aluminium disturb neuronal morphology, transport functions, cholinesterase, lysosomal and cell cycle activities. *Pathophysiology* 22: 105-115.

Akinrinade ID, Memudu AE, Ogundele OM, Ajetunmobi OI. 2015b. Interplay of glia activation and oxidative stress formation in fluoride and aluminium exposure. *Pathophysiology* 22: 39-48.

An J, Mei S, Liu A, Fu Y, Wang C. 1992. [Effect of high level of fluoride on children's intelligence]. *Chin J Control Endem Dis* 7(2): 93-94.

Aravind A, Dhanya RS, Narayan A, Sam G, Adarsh VJ, Kiran M. 2016. Effect of fluoridated water on intelligence in 10-12-year-old school children. *J Int Soc Prev Community Dent* 6(Suppl 3): S237-S242.

Arbuckle TE, Fraser WD, Fisher M, Davis K, Liang CL, Lupien N, Bastien S, Velez MP, von Dadelszen P, Hemmings DG, Wang J, Helewa M, Taback S, Sermer M, Foster W, Ross G, Fredette P, Smith G, Walker M, Shear R, Dodds L, Ettinger AS, Weber JP, D'Amour M, Legrand M, Kumarathasan P, Vincent R, Luo ZC, Platt RW, Mitchell G, Hidiroglou N, Cockell K, Villeneuve M, Rawn DF, Dabeka R, Cao XL, Becalski A, Ratnayake N, Bondy G, Jin X, Wang Z, Tittlemier S, Julien P, Avard D, Weiler H, Leblanc A, Muckle G, Boivin M, Dionne G, Ayotte P, Lanphear B, Séguin JR, Saint-Amour D, Dewailly E, Monnier P, Koren G, Ouellet E. 2013. Cohort profile: The maternal-infant research on environmental chemicals research platform. *Paediatr Perinat Epidemiol* 27(4): 415-425.

Baba NA, Raina R, Verma PK, Sultana M. 2014. Alterations in plasma and tissue acetylcholinesterase activity following repeated oral exposure of chlorpyrifos alone and in conjunction with fluoride in Wistar rats. *Proc Indian Natl Sci Acad B Biol Sci* 84: 969-972.

Bai A, Li Y, Fan Z, Li X, Li P. 2014. [Intelligence and growth development of children in coal-burning-borne arsenism and fluorosis areas: An investigation study]. *Chin J Endemiol* 33(2): 160-163.

Banala R, Nagapuri K, Mohd K, Reddy M, Karnati P. 2018. Carica Papaya leaf extract as a neuroprotective agent against behavioral and neurotransmitter changes in brain of the rat treated with sodium fluoride in pre- and post-natal periods. *Pharmacogn Mag* 14(55): 123-131.

Banala RR, Karnati PR. 2015. Vitamin A deficiency: An oxidative stress marker in sodium fluoride (NaF) induced oxidative damage in developing rat brain. *Int J Dev Neurosci* 47: 298-303.

Barberio AM, Hosein FS, Quinonez C, McLaren L. 2017a. Fluoride exposure and indicators of thyroid functioning in the Canadian population: Implications for community water fluoridation. *J Epidemiol Community Health* 71: 1019-1025.

Barberio AM, Quinonez C, Hosein FS, McLaren L. 2017b. Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation. *Can J Public Health* 108: 229-239.

Bartos M, Gumilar F, Gallegos CE, Bras C, Dominguez S, Monaco N, Esandi MDC, Bouzat C, Cancela LM, Minetti A. 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 EA, Sanchez BN, Basu N, Peterson KE, Green R, Schnaas L, Mercado-Garcia A, Hernandez-Avila M, Tellez-Rojo MM. 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 EA, Sanchez BN, Basu N, Peterson KE, Ettinger AS, Wright R, Zhang Z, Liu Y, Schnaas L, Mercado-Garcia A, Tellez-Rojo MM, Hernandez-Avila M. 2017. Prenatal fluoride exposure and cognitive outcomes in children at 4 and 6-12 years of age in Mexico. *Environ Health Perspect* 125(9): 1-12.

Begg CB, Mazumdar M. 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 50(4): 1088-1101.

Bhatnagar M, Rao P, Sushma J, Bhatnagar R. 2002. Neurotoxicity of fluoride: Neurodegeneration in hippocampus of female mice. *Indian J Exp Biol* 40: 546-554.

Bhatnagar M, Sukhwal P, Suhalka P, Jain A, Joshi C, Sharma D. 2011. Effects of fluoride in drinking water on NADPH-diaphorase neurons in the forebrain of mice: A possible mechanism of fluoride neurotoxicity. *Fluoride* 44: 195-209.

Broadbent JM, Thomson WM, Moffitt TE, Poulton R. 2015. Community water fluoridation and intelligence response. *Am J Public Health* 105: 3-4.

CDC (Centers for Disease Control and Prevention). 2013. *Community water fluoridation: Fluoridation statistics*. Atlanta, GA. Available: https://www.cdc.gov/fluoridation/statistics/2012stats.htm [accessed 19 August 2019].

Chen J, Niu Q, Xia T, Zhou G, Li P, Zhao Q, Xu C, Dong L, Zhang S, Wang A. 2018. ERK1/2-mediated disruption of BDNF-TrkB signaling causes synaptic impairment contributing to fluoride-induced developmental neurotoxicity. *Toxicology* 410: 222-230.

Chen Y. 2012. Organophosphate-induced brain damage: Mechanisms, neuropsychiatric and neurological consequences, and potential therapeutic strategies. *Neurotox* 33: 391-400.

Chen YX, Han FL, Zhoua ZL, Zhang HQ, Jiao XS, Zhang SC, Huang MC, Chang TQ, Dong YF. 1991. [Research on the intellectual development of children in high fluoride areas]. *Chin J Control Endem Dis* 6(Suppl): 99-100.

Chen YX, Han FL, Zhoua ZL, Zhang HQ, Jiao XS, Zhang SC, Huang MC, Chang TQ, Dong YF. 2008. Research on the intellectual development of children in high fluoride areas. *Fluoride* 41: 120-124.

Chinoy NJ, Narayana MV. 1992. Studies on fluorosis in Mehsana District of North Gujarat. *Proc Zool Soc* 45: 157-161.

Choi AL, Sun G, Zhang Y, Grandjean P. 2012. Developmental fluoride neurotoxicity: A systematic review and meta-analysis. *Environ Health Perspect* 120: 1362-1368.

Choi AL, Zhang Y, Sun G, Bellinger DC, Wang K, Yang XJ, Li JS, Zheng Q, Fu Y, Grandjean P. 2015. Association of lifetime exposure to fluoride and cognitive functions in Chinese children: A pilot study. *Neurotoxicol Teratol* 47: 96-101.

Chouhan S, Flora SJS. 2008. Effects of fluoride on the tissue oxidative stress and apoptosis in rats: Biochemical assays supported by IR spectroscopy data. *Toxicology* 254: 61-67.

Chouhan S, Lomash V, Flora SJ. 2010. Fluoride-induced changes in haem biosynthesis pathway, neurological variables and tissue histopathology of rats. *J Appl Toxicol* 30: 63-73.

Cochran WG. 1954. The combination of estimates from different experiments. *Biometrics* 10(1): 101-129.

Crippa A, Discacciati A, Bottai M, Spiegelman D, Orsini N. 2018. One-stage dose–response meta-analysis for aggregated data. *Stat Methods Med Res* 28(5): 1579-1596.

Cui Y, Yu J, Zhang B, Guo B, Gao T, Liu H. 2020. The relationships between thyroid-stimulating hormone and/or dopamine levels in peripheral blood and IQ in children with different urinary iodine concentrations. *Neurosci Lett* 729: 134981.

Cui Y, Zhang B, Ma J, Wang Y, Zhao L, Hou C, Yu J, Zhao Y, Zhang Z, Nie J, Gao T, Zhou G, Liu H. 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.

Das K, Mondal NK. 2016. Dental fluorosis and urinary fluoride concentration as a reflection of fluoride exposure and its impact on IQ level and BMI of children of Laxmisagar, Simlapal Block of Bankura District, W.B., India. *Environ Monit Assess* 188: 218.

Ding Y, Sun H, Han H, Wang W, Ji X, Liu X, Sun D. 2011. The relationships between low levels of urine fluoride on children's intelligence, dental fluorosis in endemic fluorosis areas in Hulunbuir, Inner Mongolia, China. *J Hazard Mater* 186: 1942-1946.

Dong YT, Wei N, Qi XL, Liu XH, Chen D, Zeng XX, Guan ZZ. 2017. Attenuating effect of vitamin E on the deficit of learning and memory of rats with chronic fluorosis: The mechanism may involve muscarinic acetylcholine receptors. *Fluoride* 50: 354-364.

Du L, Wan C, Cao X, Liu J. 1992. [The effect of fluorine on the developing human brain]. *Chin J Pathol* 21(4): 218-220.

Du L, Wan C, Cao X, Liu J. 2008. The effect of fluorine on the developing human brain. *Fluoride* 41: 327-330.

Duan J, Zhao M, Wang L, Fang D, Wang Y, Wang W. 1995. A comparative analysis of the results of multiple tests in patients with chronic industrial fluorosis. *Guizhou Med J* 18(3): 179-180.

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.

Duval S, Tweedie R. 2000a. A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *J Am Stat Assoc* 95(449): 89-98.

Duval S, Tweedie R. 2000b. Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56(2): 455-463.

Egger M, Smith G, Schneider M, Minder C, eds. 2008. *Systematic reviews in health care: meta-analysis in context*. London, UK: BMJ Publishing Group.

Egger M, Smith GD, Schneider M, Minder C. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315(7109): 629-634.

Eswar P, Nagesh L, Devaraj CG. 2011. Intelligent quotients of 12-14 year old school children in a high and low fluoride village in India. *Fluoride* 44: 168-172.

Fan Z, Dai H, Bai A, Li P, Li T, Li G. 2007. The effect of high fluoride exposure in children’s intelligence. *J Environ Health* 24(10): 802-803.

Gais S, Schonauer M. 2017. Untangling a cholinergic pathway from wakefulness to memory. *Neuron* 94(4): 696-698.

Gao Q, Liu YJ, Guan ZZ. 2009. Decreased learning and memory ability in rats with fluorosis: Increased oxidative stress and reduced cholinesterase activity in the brain. *Fluoride* 42: 277-285.

Gao Q, Liu YJ, Wu CX, Long YG, Guan ZZ. 2008a. [Effects of fluoride on learning and memory and cholinesterase activity in rat brains]. *Chin J Endemiol* 27: 128-130.

Gao Q, Liu YJ, Wu CX, Long YG, Guan ZZ. 2008b. [Level of oxidative stress in rat brains and learning and memory function of rats with chronic fluorosis]. *Chin J Endemiol* 27: 371-373.

Garman RH, Li AA, Kaufmann W, Auer RN, Bolon B. 2016. Recommended methods for brain processing and quantitative analysis in rodent developmental neurotoxicity studies. *Toxicol Pathol* 44(1): 14-42.

Ge QD, Tan Y, Luo Y, Wang WJ, Zhang H, Xie C. 2018a. MiR-132, miR-204 and BDNF-TrkB signaling pathway may be involved in spatial learning and memory impairment of the offspring rats caused by fluorine and aluminum exposure during the embryonic stage and into adulthood. *Environ Toxicol Pharmacol* 63: 60-68.

Ge Y, Chen L, Yin Z, Song X, Ruan T, Hua L, Liu J, Wang J, Ning H. 2018b. Fluoride-induced alterations of synapse-related proteins in the cerebral cortex of ICR offspring mouse brain. *Chemosphere* 201: 874-883.

Green R, Lanphear B, Hornung R, Flora D, Martinez-Mier EA, Neufeld R, Ayotte P, Muckle G, Till C. 2019. Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada. *JAMA Pediatr*: E1-E9.

Guner S, Uyar-Bozkurt S, Haznedaroglu E, Mentes A. 2016. Dental fluorosis and catalase immunoreactivity of the brain tissues in rats exposed to high fluoride pre- and postnatally. *Biol Trace Elem Res* 174: 150-157.

Guo XC, Wang RY, Cheng CF, Wei WS, Tang LM, Wang QS, Tang DX, Liu GW, He GD, Li SL. 1991. [A preliminary investigation of the IQs of 7-13 year-old children from an area with coal burning-related fluoride poisoning]. *Chin J Epidemiol* 10(2): 98-100.

Guo XC, Wang RY, Cheng CF, Wei WS, Tang LM, Wang QS, Tang DX, Liu GW, He GD, Li SL. 2008a. A preliminary investigation of the IQs of 7-13 year-old children from an area with coal burning-related fluoride poisoning. *Fluoride* 41: 125-128.

Guo ZY, He YH, Zhu QX. 2001. [Research on the neurobehavioral function of workers occupationally exposed to fluoride]. *Ind Hlth & Occup Dis* 27(6): 346-348.

Guo ZY, He YH, Zhu QX. 2008b. Research on the neurobehavioral function of workers occupationally exposed to fluoride. *Fluoride* 41: 152-155.

Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, Debeer H, Jaeschke R, Rind D, Meerpohl J, Dahm P, Schunemann HJ. 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 64(4): 383-394.

He MX, Zhang CN. 2010. [Investigation of children's intelligence quotient and dental fluorosis in drinking water-type of endemic fluorosis area in Pucheng County, Shaanxi Province before and after drinking water change]. *Chin J Endemiol* 29: 547-548.

Health Canada. 2015. *Third report on human biomonitoring of environmental chemicals in Canada - Results of the Canadian Health Measures Survey Cycle 3 (2012–2013)*. Ottawa, Ontario: Canadian Minstry of Health. Available: https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/pdf/pubs/contaminants/chms-ecms-cycle3/chms-ecms-cycle3-eng.pdf.

Higgins JP, Green S. 2011. *Cochrane handbook for systematic reviews of interventions,* In: The Cochrane Collaboration. Vol 4, New York, NY: John Wiley & Sons.

Higgins JT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. 2019. *Cochrane Handbook for systematic reviews of interventions version 6.0 (updated July 2019)*. Cochrane.

Hong FG, Cao YX, Yang D, Wang H. 2001. [Research on the effects of fluoride on child intellectual development under different environmental conditions]. *Chin Prim Health Care* 15(3): 56-57.

Hong FG, Cao YX, Yang D, Wang H. 2008. Research on the effects of fluoride on child intellectual development under different environmental conditions. *Fluoride* 41: 156-160.

Hosur MB, Puranik RS, Vanaki S, Puranik SR. 2012. Study of thyroid hormones free triiodothyronine (FT3), free thyroxine (FT4) and thyroid stimulating hormone (TSH) in subjects with dental fluorosis. *Eur J Dent* 6: 184-190.

Howard BE, Phillips J, Tandon A, Maharana A, Elmore R, Mav D, Sedykh A, Thayer K, Merrick BA, Walker V, Rooney A, Shah RR. 2020. SWIFT-Active Screener: Accelerated document screening through active learning and integrated recall estimation. *Environ Int* 138: 105623.

Jacqmin H, Commenges D, Letenneur L, Barberger-Gateau P, Dartigues JF. 1994. Components of drinking water and risk of cognitive impairment in the elderly. *Am J Epidemiol* 139: 48-57.

Jia B, Zong L, Lee JY, Lei J, Zhu Y, Xie H, Clemens JL, Feller MC, Na Q, Dong J, McLane MW, Jones-Beatty K, Burd I. 2019. Maternal supplementation of low dose fluoride alleviates adverse perinatal outcomes following exposure to intrauterine inflammation. *Scientific Reports* 9(1): 2575.

Jiang C, Zhang S, Liu H, Guan Z, Zeng Q, Zhang C, Lei R, Xia T, Wang Z, Yang L, Chen Y, Wu X, Zhang X, Cui Y, Yu L, Wang A. 2014. Low glucose utilization and neurodegenerative changes caused by sodium fluoride exposure in rat's developmental brain. *Neuromolecular Med* 16: 94-105.

Jones S, Burt BA, Petersen PE, Lennon MA. 2005. The effective use of fluorides in public health. *Bull World Health Organ* 83: 670-676.

Kang J, Cheng Y, Wu K, Lin S, He G, Jin Y. 2011. [Effect of exposure to fluoride and arsenic in drinking water of Hangjinhouqi on children's intelligence]. *Chin School Health*: 679-681.

Karimzade S, Aghaei M, Mahvi AH. 2014. Investigation of intelligence quotient in 9-12 year-old children exposed to high- and low-drinking water fluoride in West Azerbaijan Province, Iran. *Fluoride* 47: 9-14.

Khan AM, Raina R, Dubey N, Verma PK. 2017. Effect of deltamethrin and fluoride co-exposure on the brain antioxidant status and cholinesterase activity in Wistar rats. *Drug Chem Toxicol* 41(2): 1-5.

Khan SA, Singh RK, Navit S, Chadha D, Johri N, Navit P, Sharma A, Bahuguna R. 2015. Relationship between dental fluorosis and intelligence quotient of school going children in and around Lucknow District: A cross-sectional study. *J Clin Diagn Res* 9(11): 10-15.

Khandare AL, Gourineni SR, Validandi V. 2017. Dental fluorosis, nutritional status, kidney damage, and thyroid function along with bone metabolic indicators in school-going children living in fluoride-affected hilly areas of Doda District, Jammu and Kashmir, India. *Environ Monit Assess* 189: 579.

Khandare AL, Validandi V, Gourineni SR, Gopalan V, Nagalla B. 2018. Dose-dependent effect of fluoride on clinical and subclinical indices of fluorosis in school going children and its mitigation by supply of safe drinking water for 5 years: An Indian study. *Environ Monit Assess* 190: 110.

Kheradpisheh Z, Mahvi AH, Mirzaei M, Mokhtari M, Azizi R, Fallahzadeh H, Ehrampoush MH. 2018a. Correlation between drinking water fluoride and TSH hormone by ANNs and ANFIS. *J Environ Health Sci Eng* 16(1): 11-18.

Kheradpisheh Z, Mirzaei M, Mahvi AH, Mokhtari M, Azizi R, Fallahzadeh H, Ehrampoush MH. 2018b. Impact of drinking water fluoride on human thyroid hormones: A case-control study. *Sci Rep* 8: 2674.

Kinawy AA, Al-Eidan AA. 2018. Impact of prenatal and postnatal treatment of sodium fluoride and aluminum chloride on some hormonal and sensorimotor aspects in rats. *Biol Trace Elem Res* 186: 1-8.

Kumar V, Chahar P, Kajjari S, Rahman F, Bansal DK, Kapadia JM. 2018. Fluoride, thyroid hormone derangements and its correlation with tooth eruption pattern among the pediatric population from endemic and non-endemic fluorosis areas. *J Contemp Dent Pract* 19(12): 1512-1516.

Kundu H, Basavaraj P, Singla A, Gupta R, Singh K, Jain S. 2015. Effect of fluoride in drinking water on children's intelligence in high and low fluoride areas of Delhi. *J Indian Assoc Public Health Dent* 13(2): 116-121.

Li F, Chen X, Huang R, Xie Y. 2009. The impact of endemic fluorosis caused by the burning of coal on the development of intelligence in children. *J Environ Health* 26(4): 838-840.

Li J, Yao L, Shao QL, Wu CY. 2004. [Effects of high fluoride level on neonatal neurobehavioral development]. *Chin J Endemiol* 23(5): 463-465.

Li J, Yao L, Shao QL, Wu CY. 2008a. Effects of high fluoride level on neonatal neurobehavioral development. *Fluoride* 41: 165-170.

Li M, Gao Y, Cui J, Li Y, Li B, Liu Y, Sun J, Liu X, Liu H, Zhao L, Sun D. 2016. Cognitive impairment and risk factors in elderly people living in fluorosis areas in China. *Biol Trace Elem Res* 172: 53-60.

Li X, Hou G, Yu B, Yuan C, Liu Y, L Z. 2010. [Investigation and analysis of children's IQ and dental fluorosis in high fluoride area]. *Chin J Pest Control* 26(3): 230-231.

Li X, Zhang J, Niu R, Manthari RK, Yang K, Wang J. 2019. Effect of fluoride exposure on anxiety- and depression-like behavior in mouse. *Chemosphere* 215: 454-460.

Li XS, Zhi JL, Gao RO. 1995. Effect of fluoride exposure on the intelligence of children. *Fluoride* 28: 189-192.

Li Y, Li X, Wei S. 1994. [Effects of high fluoride intake on child mental work capacity: Preliminary investigation into the mechanisms involved]. *J West Chin University Med Sci* 25(2): 188-191.

Li Y, Li X, Wei S. 2008b. Effects of high fluoride intake on child mental work capacity: Preliminary investigation into the mechanisms involved. *Fluoride* 41: 331-335.

Li YP, Jing XY, Chen D, Lin L, Wang ZJ. 2003. [Effects of endemic fluoride poisoning on the intellectual development of children in Baotou]. *Chin J Public Health Manag* 19(4): 337-338.

Li YP, Jing XY, Chen D, Lin L, Wang ZJ. 2008c. Effects of endemic fluoride poisoning on the intellectual development of children in Baotou. *Fluoride* 41: 161-164.

Lin F, Ai H, Zhao H, Lin J, Jhiang J, Maimaiti. 1991. [High fluoride and low iodine environment and subclinical cretinism in Xinjiang]. *Endem Dis Bull* 6(2): 62-67.

Liu S, Lu Y, Sun Z, Wu L, Wang X, Yan S. 2000. [Report on the intellectual ability of children living in high-fluoride water areas]. *Chin J Control Endem Dis* 15(4): 231-232.

Liu SL, Lu Y, Sun ZR, Wu L, Lu WL, Wang XW, Yan S. 2008. Report on the intellectual ability of children living in high-fluoride water areas. *Fluoride* 41: 144-147.

Liu YJ, Gao Q, Wu CX, Guan ZZ. 2010. Alterations of nAChRs and ERK1/2 in the brains of rats with chronic fluorosis and their connections with the decreased capacity of learning and memory. *Toxicol Lett* 192: 324-329.

Lou DD, Guan ZZ, Liu YJ, Liu YF, Zhang KL, Pan JG, Pei JJ. 2013. The influence of chronic fluorosis on mitochondrial dynamics morphology and distribution in cortical neurons of the rat brain. *Arch Toxicol* 87: 449-457.

Lu F, Zhang Y, Trivedi A, Jiang X, Chandra D, Zheng J, Nakano Y, Uyghurturk DA, Jalai R, Guner S, Mentes A, DenBesten PK. 2019. Fluoride related changes in behavioral outcomes may relate to increased serotonin. *Physiol Behav* 206: 76-83.

Lu Y, Sun ZR, Wu LN, Wang X, Lu W, Liu SS. 2000. Effect of high-fluoride water on intelligence in children. *Fluoride* 33: 74-78.

Ma Q, Huang H, Sun L, Zhou T, Zhu J, Cheng X, Duan L, Li Z, Cui L, Ba Y. 2017. Gene-environment interaction: Does fluoride influence the reproductive hormones in male farmers modified by ERα gene polymorphisms? *Chemosphere* 188: 525-531.

Malin AJ, 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 AJ, 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.

Manusha S, Sudhakar K, Reddy KP. 2019. Protective effects of allium sativum extract against sodium fluoride induced neurotoxicity. *Int J Pharm Sci Res* 10(2): 625-633.

McPherson CA, Zhang G, Gilliam R, Brar SS, Wilson R, Brix A, Picut C, Harry GJ. 2018. An evaluation of neurotoxicity following fluoride exposure from gestational through adult ages in Long-Evans hooded rats. *Neurotoxicol Res*: 1-18.

Mesram N, Nagapuri K, Banala RR, Nalagoni CR, Karnati PR. 2016. Quercetin treatment against NaF induced oxidative stress related neuronal and learning changes in developing rats. *J King Saud Univ Sci* 29: 221-229.

Michael M, Barot VV, Chinoy NJ. 1996. Investigations of soft tissue functions in fluorotic individuals of North Gujarat. *Fluoride* 29: 63-71.

Miller K, Howard B, Phillips J, Shah M, Mav D, Thayer K, Shah R. 2016. SWIFT-Active screener: Reducing literature screening effort through machine learning for systematic reviews, Cochrane Colloquium Seoul, Seoul, Korea.

Moher D, Liberati A, Tetzlaff J, Altman D. 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 6(7): e1000097.

Mondal D, Dutta G, Gupta S. 2016. Inferring the fluoride hydrogeochemistry and effect of consuming fluoride-contaminated drinking water on human health in some endemic areas of Birbhum District, West Bengal. *Environ Geochem Health* 38: 557-576.

Morgan L, Allred E, Tavares M, Bellinger D, Needleman H. 1998. Investigation of the possible associations between fluorosis, fluoride exposure, and childhood behavior problems. *Pediatr Dent* 20: 244-252.

Müller S, Scealy JL, Welsh AH. 2013. Model selection in linear mixed models. *Statist Sci* 28(2): 135-167.

Mustafa DE, Younis UM, Elhag SAA. 2018. The relationship between the fluoride levels in drinking water and the schooling performance of children in rural areas of Khartoum State, Sudan. *Fluoride* 51(2): 102-113.

Nagarajappa R, Pujara P, Sharda AJ, Asawa K, Tak M, Aapaliya P, Bhanushali N. 2013. Comparative assessment of intelligence quotient among children living in high and low fluoride areas of Kutch, India: A pilot study. *Iran J Public Health* 42: 813-818.

Nageshwar M, Sudhakar K, Reddy KP. 2018. Quercetin ameliorates oxidative stress, neural damage of brain and behavioral impairment of rat with fluoride exposure. *Int J Pharm Sci Res* 9(8): 3247-3256.

Nageshwar M, Sudhakar K, Reddy NCC, Reddy KP. 2017. Neuroprotective effects of curcumin on sodium fluoride induced behavioural and enzymatic changes in brain and muscles of rat. *J Environ Biol* 38: 675-681.

NASEM (National Academies of Sciences, Engineering and Medicine). 2020. *Review of the draft NTP monograph: Systematic review of fluoride exposure and neurodevelopmental and cognitive health effects*. Washington, DC: The National Academies Press. Available: https://doi.org/10.17226/25715.

NIOSH (National Institute for Occupational Safety and Health). 1984. *Fluoride in urine*. In: Manual of Analytical Methods Vol 11. Method 8308. Washington, DC: US Department of Health and Human Services: 1-3.

Niu Q, Chen J, Xia T, Li P, Zhou G, Xu C, Zhao Q, Dong L, Zhang S, Wang A. 2018. Excessive ER stress and the resulting autophagic flux dysfunction contribute to fluoride-induced neurotoxicity. *Environ Pollut* 233: 889-899.

Nkpaa KW, Onyeso GI. 2018. Rutin attenuates neurobehavioral deficits, oxidative stress, neuro-inflammation and apoptosis in fluoride treated rats. *Neurosci Lett* 682: 92-99.

NRC (National Research Council). 2006. *Committee on fluoride in drinking water, board on environmental studies and toxicology. Fluoride in drinking water: A scientific review of EPA's standards.* Available: http://www.nap.edu/catalog/11571/fluoride-in-drinking-water-a-scientific-review-of-epas-standards [accessed 19 August 2019].

NTP (National Toxicology Program). 2015. *Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration*. Research Triangle Park, NC. Available: http://ntp.niehs.nih.gov/go/38673 [accessed 19 August 2019].

NTP (National Toxicology Program). 2016. *Systematic literature review on the effects of fluoride on learning and memory in animal studies*. NTP Research Report 1. Research Triangle Park, NC. Available: https://ntp.niehs.nih.gov/ntp/results/pubs/rr/reports/01fluoride_508.pdf [accessed 19 August 2019].

NTP (National Toxicology Program). 2019. *Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration*. Research Triangle Park, NC. Available: https://ntp.niehs.nih.gov/go/systematic_review [accessed 19 August 2019].

OEHHA (California Office of Environmental Health Hazard Assessment). 2011. *Meeting synopsis and slide presentations: carcinogen identification committee meeting held on October 12, 2011*. Available: http://oehha.ca.gov/prop65/public_meetings/cic101211synop.html [accessed 19 August 2019].

Peckham S, Lowery D, Spencer S. 2015. Are fluoride levels in drinking water associated with hypothyroidism prevalence in England? A large observational study of GP practice data and fluoride levels in drinking water. *J Epidemiol Community Health* 69: 619-624.

Podgorski J, Berg M. 2020. Global threat of arsenic in groundwater. *Science* 368(6493): 845-850.

Poureslami HR, Horri A, Garrusi B. 2011. A comparative study of the IQ of children age 7-9 in a high and a low fluoride water city in Iran. *Fluoride* 44: 163-167.

Pulungan ZS, Sofro ZM, Partadiredja G. 2016. Sodium fluoride does not affect the working memory and number of pyramidal cells in rat medial prefrontal cortex. *Anat Sci Int* 93(1): 128-138.

Qin LS, Huo SY, Chen RL, Chang YZ, Zhao MY. 1990. [Using the Raven's Standard Progressive Matrices to determine the effects of the level of fluoride in drinking water on the intellectual ability of school-age children]. *J Control Endem Dis* 5(4): 203-204.

Qin LS, Huo SY, Chen RL, Chang YZ, Zhao MY. 2008. Using the Raven's Standard Progressive Matrices to determine the effects of the level of fluoride in drinking water on the intellectual ability of school-age children. *Fluoride* 41: 115-119.

Raju S, Sivanesan SK, Gudemalla K. 2019. Cognitive enhancement effect of Ginkgo biloba extract on memory and learning impairments induced by fluoride neurotoxicity. *Int J Res Pharm Sci* 10(1): 129-134.

Razdan P, Patthi B, Kumar JK, Agnihotri N, Chaudhari P, Prasad M. 2017. Effect of fluoride concentration in drinking water on intelligence quotient of 12-14-year-old children in Mathura District: A cross-sectional study. *J Int Soc Prev Community Dent* 7: 252-258.

Reddy YP, Tiwari SK, Shaik AP, Alsaeed A, Sultana A, Reddy PK. 2014. Effect of sodium fluoride on neuroimmunological parameters, oxidative stress and antioxidative defenses. *Toxicol Mech Methods* 24: 31-36.

Ren D, Li K, Liu D. 1989. [A study of the intellectual ability of 8-14 year-old children in high fluoride, low iodine areas]. *Chin J Control Endem Dis* 4(4): 251.

Ren D, Li K, Liu D. 2008. A study of the intellectual ability of 8-14 year-old children in high fluoride, low iodine areas. *Fluoride* 41: 319-320.

Riddell JK, Malin AJ, 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: 105190.

Rocha-Amador D, Navarro M, Trejo-Acevedo A, Carrizales L, Perez-Maldonado I, Diaz-Barriga F, Calderon J. 2009. Use of the Rey-Osterrieth Complex Figure Test for neurotoxicity evaluation of mixtures in children. *Neurotox* 30: 1149-1154.

Rocha-Amador D, Navarro ME, Carrizales L, Morales R, Calderon J. 2007. Decreased intelligence in children and exposure to fluoride and arsenic in drinking water. *Cad Saude Publica* 23(Suppl 4): S579-587.

Rooney AA, Boyles AL, Wolfe MS, Bucher JR, Thayer KA. 2014. Systematic review and evidence integration for literature-based environmental health science assessments. *Environ Health Perspect* 122(7): 711-718.

Rooney AA, Cooper GS, Jahnke GD, Lam J, Morgan RL, Ratcliffe JM, Kraft AD, Schünemann HJ, Schwingl P, Walker TD, Thayer KA, Lunn RM. 2016. How credible are the study results? Evaluating and applying internal validity tools to literature-based assessments of environmental health hazards. *Environ Int* 92-93: 617-629.

Rotton J, Tikofsky RS, Feldman HT. 1982. Behavioral effects of chemicals in drinking water. *J Appl Psychol* 67: 230-238.

Russ TC, Killin LOJ, Hannah J, Batty GD, Deary IJ, Starr JM. 2019. Aluminium and fluoride in drinking water in relation to later dementia risk. *Brit J Psychol* 14: 1-6.

Saxena S, Sahay A, Goel P. 2012. Effect of fluoride exposure on the intelligence of school children in Madhya Pradesh, India. *J Neurosci Rural Pract* 3: 144-149.

SCHER (Scientific Committee on Health and Environmental Risks). 2011. *Critical review of any new evidence on the hazard profile, health effects, and human exposure to fluoride and the fluoridating agents of drinking water*. European Commission Directorate-General for Health and Consumers Scientific Committees. Available: http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_139.pdf [accessed 19 August 2019].

Sebastian ST, Sunitha S. 2015. A cross-sectional study to assess the intelligence quotient (IQ) of school going children aged 10-12 years in villages of Mysore District, India with different fluoride levels. *J Indian Soc Pedod Prev Dent* 33: 307-311.

Seraj B, Shahrabi M, Falahzade M, Falahzade F, Akhondi N. 2006. Effect of high fluoride concentration in drinking water on children's intelligence. *J Dent Med* 19(2): 80-86.

Seraj B, Shahrabi M, Shadfar M, Ahmadi R, Fallahzadeh M, Eslamlu HF, Kharazifard MJ. 2012. Effect of high water fluoride concentration on the intellectual development of children in Makoo, Iran. *J Dent* 9: 221-229.

Shalini B, Sharma JD. 2015. Beneficial effects of emblica officinalis on fluoride-induced toxicity on brain biochemical indexes and learning-memory in rats. *Toxicol Int* 22: 35-39.

Shan KR, Qi XL, Long YG, Nordberg A, Guan ZZ. 2004. Decreased nicotinic receptors in PC12 cells and rat brains influenced by fluoride toxicity: A mechanism relating to a damage at the level in post-transcription of the receptor genes. *Toxicology* 200: 169-177.

Shannon FT, Fergusson DM, Horwood LJ. 1986. Exposure to fluoridated public water supplies and child health and behaviour. *New Zeal Med J* 99: 416-418.

Shao Q. 2003. [Study of cognitive function impairment caused by chronic fluorosis]. *Chin J Endemiol* 22(4): 336-338.

Sharma C, Suhalka P, Bhatnagar M. 2018. Curcumin and resveratrol rescue cortical-hippocampal system from chronic fluoride-induced neurodegeneration and enhance memory retrieval. *Int J Neurosci*: 1-15.

Sharma JD, Sohu D, Jain P. 2009. Prevalence of neurological manifestations in a human population exposed to fluoride in drinking water. *Fluoride* 42: 127-132.

Shivaprakash PK, Ohri K, Noorani H. 2011. Relation between dental fluorosis and intelligence quotient in school children of Bagalkot District. *J Indian Soc Pedod Prev Dent* 29: 117-120.

Singh N, Verma KG, Verma P, Sidhu GK, Sachdeva S. 2014. A comparative study of fluoride ingestion levels, serum thyroid hormone & TSH level derangements, dental fluorosis status among school children from endemic and non-endemic fluorosis areas. *Springerplus* 3: 7.

Singh V, Singh C, Sandeep T, Sandeep K, Vikas G, Mukesh T, Anurag T. 2013. A correlation between serum vitamin, acetylcholinesterase activity and IQ in children with excessive endemic fluoride exposure in Rajasthan, India. *Int Res J Medical Sci* 1(3): 12-16.

Soto-Barreras U, Escalante-Villalobos KY, Holguin-Loya B, Perez-Aguirre B, Nevarez-Rascon A, Martinez-Martinez RE, Loyola-Rodriguez JP. 2019. Effect of fluoride in drinking water on dental caries and IQ in children. *Fluoride* 52: 474-482.

Still CN, Kelley P. 1980. The incidence of primary degenerative dementia vs. water fluoride content in South Carolina, USA. *Neurotox* 1: 125-132.

Sudhakar K, Nageshwar M, Pratap Reddy K. 2017. Seed extract of Abelmoschus moschatus medik reverses NAF-induced behavioral changes through neurodegeneration and oxidative stress in brain of rat. *Asian J Pharm Clin Res* 10: 165-171.

Sudhakar K, Nageshwar M, Reddy KP. 2018a. Abelmoschus moschatus extract reverses altered pain and neurohistology of a rat with developmental exposure of fluoride. *J Appl Pharm Sci* 8(6): 94-104.

Sudhakar K, Nageshwar M, Reddy KP. 2018b. Protective effect of okra, Abelmoschus moschatus seed extract on developing brain of rats during pre- and post-natal fluoride exposure. *Int J Pharm Sci Res* 9: 1519-1528.

Sudhakar K, Reddy KP. 2018. Protective effects of Abelmoschus moschatus seed extract on neurotransmitter system of developing brain of Wistar rats with gestational and post-natal exposure of sodium fluoride. *Int J Green Pharm* 12: S131-S139.

Sudhir KM, Chandu GN, Prashant GM, Subba Reddy VV. 2009. Effect of fluoride exposure on intelligence quotient (IQ) among 13-15 year old school children of known endemic area of fluorosis, Nalgonda District, Andhra Pradesh. *J Indian Assoc Public Health Dent* 2009(13): 88-94.

Sun M, Li S, Wang Y, Li F. 1991. [Measurement of intelligence by drawing test among the children in the endemic area of Al-F combined toxicosis]. *J Guiyang Med Coll* 16(3): 204-206.

Sun Z, Zhang Y, Xue X, Niu R, Wang J. 2018. Maternal fluoride exposure during gestation and lactation decreased learning and memory ability, and glutamate receptor mRNA expressions of mouse pups. *Hum Exp Toxicol* 37: 87-93.

Sun ZR, Liu FZ, Wu LN, Lu Y, Yu DK. 2000. [Effects of high fluoride drinking water on the cerebral function of mice]. *Chin J Epidemiol* 19: 262-263.

Sun ZR, Liu FZ, Wu LN, Lu Y, Yu DK. 2008. Effects of high fluoride drinking water on the cerebral function of mice. *Fluoride* 41: 148-151.

Susheela AK, Bhatnagar M, Vig K, Mondal NK. 2005. Excess fluoride ingestion and thyroid hormone derangements in children living in Delhi, India. *Fluoride* 38: 98-108.

Till C, Green R, Flora D, Hornung R, Martinez-Mier EA, Blazer M, Farmus L, Ayotte P, Muckle G, Lanphear B. 2020. Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. *Environ Int* 134: 105315.

Till C, Green R, Grundy JG, Hornung R, Neufeld R, Martinez-Mier EA, Ayotte P, Muckle G, Lanphear B. 2018. Community water fluoridation and urinary fluoride concentrations in a national sample of pregnant women in Canada. *Environ Health Perspect* 126(10): 107001.

Trivedi M, Sangai N, Patel R, Payak M, Vyas S. 2012. Assessment of groundwater quality with special reference to fluoride and its impact on IQ of schoolchildren in six villages of the Mundra Region, Kachchh, Gujurat, India. *Fluoride* 45(4): 377-383.

Trivedi MH, Verma RJ, Chinoy NJ, Patel RS, Sathawara NG. 2007. Effect of high fluoride water on intelligence of school children in India. *Fluoride* 40: 178-183.

Tsunoda M, Aizawa Y, Nakano K, Liu Y, Horiuchi T, Itai K, Tsunoda H. 2005. Changes in fluoride levels in the liver, kidney, and brain and in neurotransmitters of mice after subacute administration of fluoride. *Fluoride* 38: 284-292.

US DHHS (U.S. Department of Health and Human Services). 2015. *U.S. Public Health Service recommendation for fluoride concentration in drinking water for the prevention of dental caries*. 318-331. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4547570/ [accessed 19 August 2019].

US EPA (U.S. Environmental Protection Agency). 2010. *Fluoride: Exposure and relative source contribution analysis*. 820-R-10-015. Washington, DC. Available: http://www.epa.gov/dwstandardsregulations/fluoride-risk-assessment-and-relative-source-contribution [accessed 19 August 2019].

Valdez Jimenez L, Lopez Guzman OD, Cervantes Flores M, Costilla-Salazar R, Calderon Hernandez J, Alcaraz Contreras Y, Rocha-Amador DO. 2017. In utero exposure to fluoride and cognitive development delay in infants. *Neurotox* 59: 65-70.

Villa A, Anabalon M, Zohouri V, Maguire A, Franco AM, Rugg-Gunn A. 2010. Relationships between fluoride intake, urinary fluoride excretion and fluoride retention in children and adults: An analysis of available data. *Caries Res* 44(1): 60-68.

Wang A, Duan L, Huang H, Ma J, Zhang Y, Ma Q, Guo Y, Li Z, Cheng X, Zhu J, Zhou G, Ba Y. 2020a. Association between fluoride exposure and behavioural outcomes of school-age children: A pilot study in China. *Int J Environ Health Res*: 1-10.

Wang G, Gao M, Zhang M, Yang M, Xiang Q. 2012. Correlation between total fluoride intake and children's IQ. *J Southeast Univ Med Ed*: 743-746.

Wang G, Yang D, Jia F, Wang H. 1996. [A study of the IQ levels of four- to seven-year-old children in high fluoride areas]. *Endem Dis Bull* 11(1): 60-62.

Wang G, Yang D, Jia F, Wang H. 2008b. A study of the IQ levels of four- to seven-year-old children in high fluoride areas. *Fluoride* 41: 340-343.

Wang J, Zhang Y, Guo Z, Li R, Xue X, Sun Z, Niu R. 2018. Effects of perinatal fluoride exposure on the expressions of miR-124 and miR-132 in hippocampus of mouse pups. *Chemosphere* 197: 117-122.

Wang M, Liu L, Li H, Li Y, Liu H, Hou C, Zeng Q, Li P, Zhao Q, Dong L, Zhou G, Yu X, Liu L, Guan Q, Zhang S, Wang A. 2020b. Thyroid function, intelligence, and low-moderate fluoride exposure among Chinese school-age children. *Environ Int* 134: 105229.

Wang SX, Wang ZH, Cheng XT, Li J, Sang ZP, Zhang XD, Han LL, Qiao XY, Wu ZM, Wang ZQ. 2005. [Investigation and evaluation on intelligence and growth of children in endemic fluorosis and arsenism areas]. *Chin J Endemiol* 24: 179-182.

Wang SX, Wang ZH, Cheng XT, Li J, Sang ZP, Zhang XD, Han LL, Qiao XY, Wu ZM, Wang ZQ. 2007. Arsenic and fluoride exposure in drinking water: Children's IQ and growth in Shanyin County, Shanxi Province, China. *Environ Health Perspect* 115: 643-647.

Wang SY, Zhang HX, Fan W, Fang SJ, Kang PP, Chen XG, Yu MJ. 2005b. [The effects of endemic fluoride poisoning caused by coal burning on the physical development and intelligence of children]. *J Appl Clin Ped* 20(9): 897-899.

Wang SY, Zhang HX, Fan W, Fang SJ, Kang PP, Chen XG, Yu MJ. 2008a. The effects of endemic fluoride poisoning caused by coal burning on the physical development and intelligence of children. *Fluoride* 41: 344-348.

Wang X, Wang L, Hu P, Guo X, Luo X. 2001. [Effects of high iodine and high fluorine on children's intelligence and thyroid function]. *Chin J Endemiol* 20(4): 288-290.

Wang Z, Wang S, Zhang X, Li J, Zheng X, Hu C. 2006. Investigation of children's growth and development under long-term fluoride exposure. *Chin J Control Endem Dis* 21(4): 239-241.

Watanabe M, Kono K, Orita Y, Dote T, Usuda K, Takahashi Y, Yoshida Y. 1995. Influence of dietary fluoride intake on urinary fluoride concentration and evaluation of corrected levels in spot urine. *Fluoride* 28(2): 61-70.

Wei N, Li Y, Deng J, Xu S, Guan Z. 2014. [The effects of comprehensive control measures on intelligence of school-age children in coal-burning-borne endemic fluorosis areas]. *Chin J Endemiol* 33(3): 320-322.

WHO (World Health Organization). 2008. *Guidelines for drinking-water quality [electronic resource]: Incorporating 1st and 2nd addenda*. Third Edition. Vol. 1. Geneva, Switzerland. Available: https://apps.who.int/iris/bitstream/handle/10665/204411/9789241547611_eng.pdf?sequence=1&isAllowed=y.

WHO (World Health Organization). 2011. *Guidelines for drinking-water quality*. Fourth edition. Available: https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151_eng.pdf?sequence=1&isAllowed=y&ua=1.

Xiang Q, Liang Y, Chen B, Chen L. 2011. Analysis of children's serum fluoride levels in relation to intelligence scores in a high and low fluoride water village in China. *Fluoride* 44: 191-194.

Xiang Q, Liang Y, Chen L, Wang C, Chen B, Chen X, Zhou M. 2003a. Effect of fluoride in drinking water on children's intelligence. *Fluoride* 36: 84-94.

Xiang Q, Wang Y, Yang M, Zhang M, Xu Y. 2013. Level of fluoride and arsenic in household shallow well water in Wamiao and Xinhuai villages in Jiangsu province, China. *Fluoride* 46: 192-197.

Xiang QY, Liang YX, Zhou MS, Zang HB. 2003b. Blood lead of children in Wamiao-Xinhuai intelligence study. *Fluoride* 36: 198-199.

Xu Y, Lu C, Zhang X. 1994. [The effect of fluorine on the level of intelligence in children]. *Endem Dis Bull* 9(2): 83-84.

Yang L, Jin P, Wang X, Zhou Q, Lin X, Xi S. 2018. Fluoride activates microglia, secretes inflammatory factors and influences synaptic neuron plasticity in the hippocampus of rats. *Neurotox* 69: 108-120.

Yang Y, Wang X, Guo X, Hu P. 1994. [The effects of high levels of fluoride and iodine on child intellectual ability and the metabolism of fluoride and iodine]. *Chin J Epidemiol* 15(4): 296-298.

Yang Y, Wang X, Guo X, Hu P. 2008. The effects of high levels of fluoride and iodine on child intellectual ability and the metabolism of fluoride and iodine. *Fluoride* 41: 336-339.

Yao L, Zhou J, Wang S, Cui K, Lin F. 1996. Analysis on TSH and intelligence level of children with dental Fluorosis in a high fluoride area. *Lit Inf Prev Med* 2(1): 26-27.

Yao Y. 1997. Comparable analysis on the physical and mental development of children in endemic fluorosis area with water improvement and without water improvement. *Lit Inf Prev Med* 3(1): 42-43.

Yasmin S, Ranjan S, D'Souza D. 2013. Effect of excess fluoride ingestion on human thyroid function in Gaya Region, Bihar, India. *Toxicol Environ Chem* 95: 1235-1243.

Yazdi SM, Sharifian A, Dehghani-Beshne M, Momeni VR, Aminian O. 2011. Effects of fluoride on psychomotor performance and memory of aluminum potroom workers. *Fluoride* 44: 158-162.

Yu X, Chen J, Li Y, Liu H, Hou C, Zeng Q, Cui Y, Zhao L, Li P, Zhou Z, Pang S, Tang S, Tian K, Zhao Q, Dong L, Xu C, Zhang X, Zhang S, Liu L, Wang A. 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.

Yu YN, Yang WX, Dong Z, Wan CW, Zhang JT, Liu JL, Xiao KQ, Huang YS, Lu BF. 1996. [Neurotransmitter and receptor changes in the brains of fetuses from areas of endemic fluorosis]. *Chin J Endemiol* 15(5): 257-259.

Yu YN, Yang WX, Dong Z, Wan CW, Zhang JT, Liu JL, Xiao KQ, Huang YS, Lu BF. 2008. Neurotransmitter and receptor changes in the brains of fetuses from areas of endemic fluorosis. *Fluoride* 41: 134-138.

Yuan J, Li Q, Niu R, Wang J. 2019. Fluoride exposure decreased learning ability and the expressions of the insulin receptor in male mouse hippocampus and olfactory bulb. *Chemosphere* 224: 71-76.

Zhang J, Yao H, Chen Y. 1998. [The effect of high levels of arsenic and fluoride on the development of children's intelligence]. *Chin J Public Health* 17(2): 119.

Zhang KL, Lou DD, Guan ZZ. 2015a. Activation of the AGE/RAGE system in the brains of rats and in SH-SY5Y cells exposed to high level of fluoride might connect to oxidative stress. *Neurotoxicol Teratol* 48: 49-55.

Zhang S, Zhang X, Liu H, Qu W, Guan Z, Zeng Q, Jiang C, Gao H, Zhang C, Lei R, Xia T, Wang Z, Yang L, Chen Y, Wu X, Cui Y, Yu L, Wang A. 2015b. 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: 238-245.

Zhao LB, Liang GH, Zhang DN, Wu XR. 1996. Effect of a high fluoride water supply on children's intelligence. *Fluoride* 29: 190-192.

Zhao Q, Niu Q, Chen J, Xia T, Zhou G, Li P, Dong L, Xu C, Tian Z, Luo C, Liu L, Zhang S, Wang A. 2019. Roles of mitochondrial fission inhibition in developmental fluoride neurotoxicity: Mechanisms of action in vitro and associations with cognition in rats and children. *Arch Toxicol* 93(3): 709-726.

Zhao Q, Tian Z, Zhou G, Niu Q, Chen J, Li P, Dong L, Xia T, Zhang S, Wang A. 2020. SIRT1-dependent mitochondrial biogenesis supports therapeutic effects of resveratrol against neurodevelopment damage by fluoride. *Theranostics* 10: 4822-4838.

Zheng X, Sun Y, Ke L, Ouyang W, Zhang Z. 2016. Molecular mechanism of brain impairment caused by drinking-acquired fluorosis and selenium intervention. *Environ Toxicol Pharmacol* 43: 134-139.

Zhou G, Tang S, Yang L, Niu Q, Chen J, Xia T, Wang S, Wang M, Zhao Q, Liu L, Li P, Dong L, Yang K, Zhang S, Wang A. 2019. Effects of long-term fluoride exposure on cognitive ability and the underlying mechanisms: Role of autophagy and its association with apoptosis. *Toxicol Appl Pharmacol* 378: 114608.

Zhou T, Duan L-J, Ding Z, Yang R-P, Li S-H, Xi Y, Cheng X-M, Hou J-X, Wen S-B, Chen J, Cui L-X, Ba Y. 2012. Environmental fluoride exposure and reproductive hormones in male living in endemic fluorosis villages in China. *Life Sci J* 9(4): 1-7.

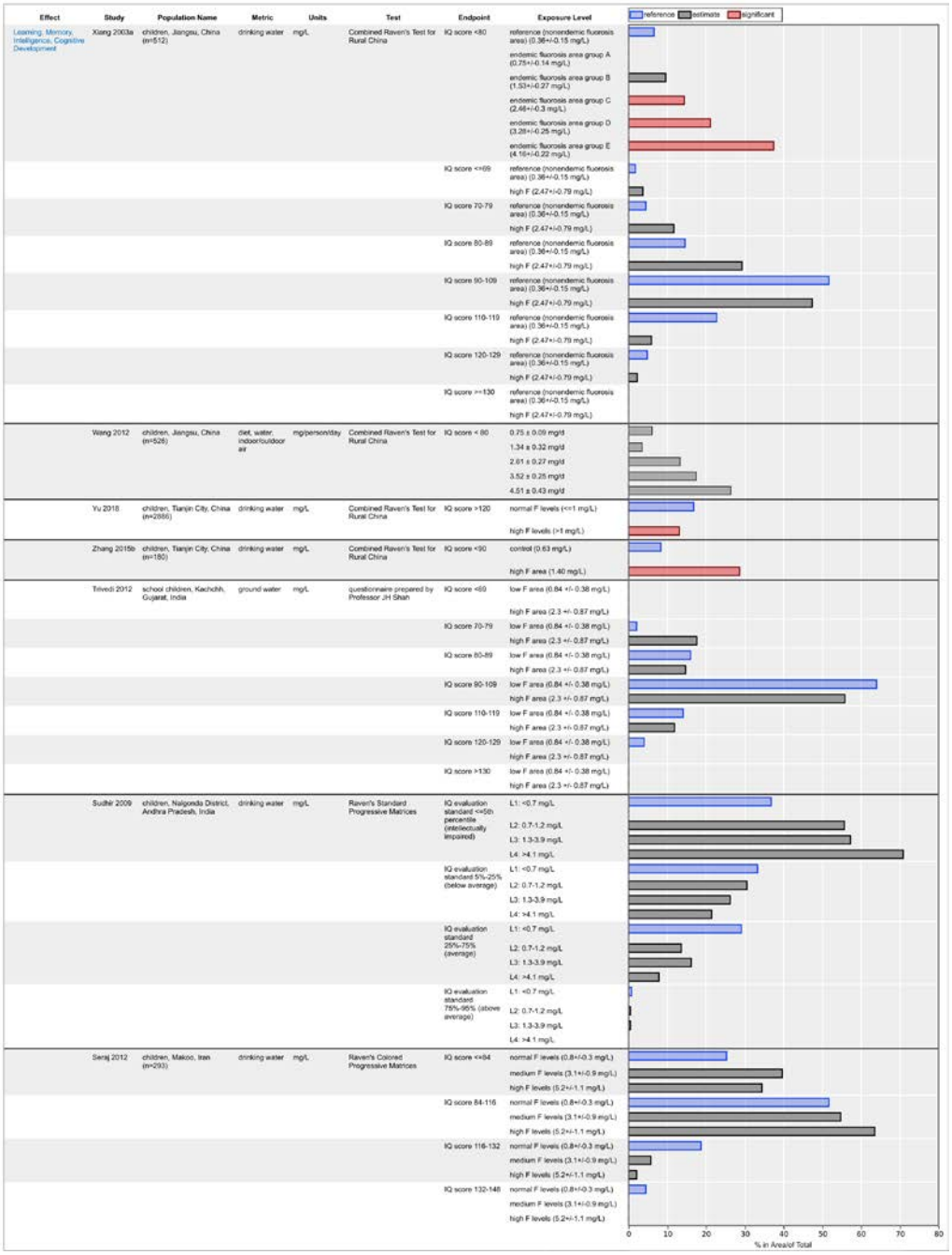
Zhu YP, Xi SH, Li MY, Ding TT, Liu N, Cao FY, Zeng Y, Liu XJ, Tong JW, Jiang SF. 2017. Fluoride and arsenic exposure affects spatial memory and activates the ERK/CREB signaling pathway in offspring rats. *Neurotox* 59: 56-64.

Zohouri FV, Swinbank CM, Maguire A, Moynihan PJ. 2006. Is the fluoride/creatinine ratio of a spot urine sample indicative of 24-h urinary fluoride? *Community Dent Oral Epidemiol* 34(2): 130-138.

DATA FIGURES

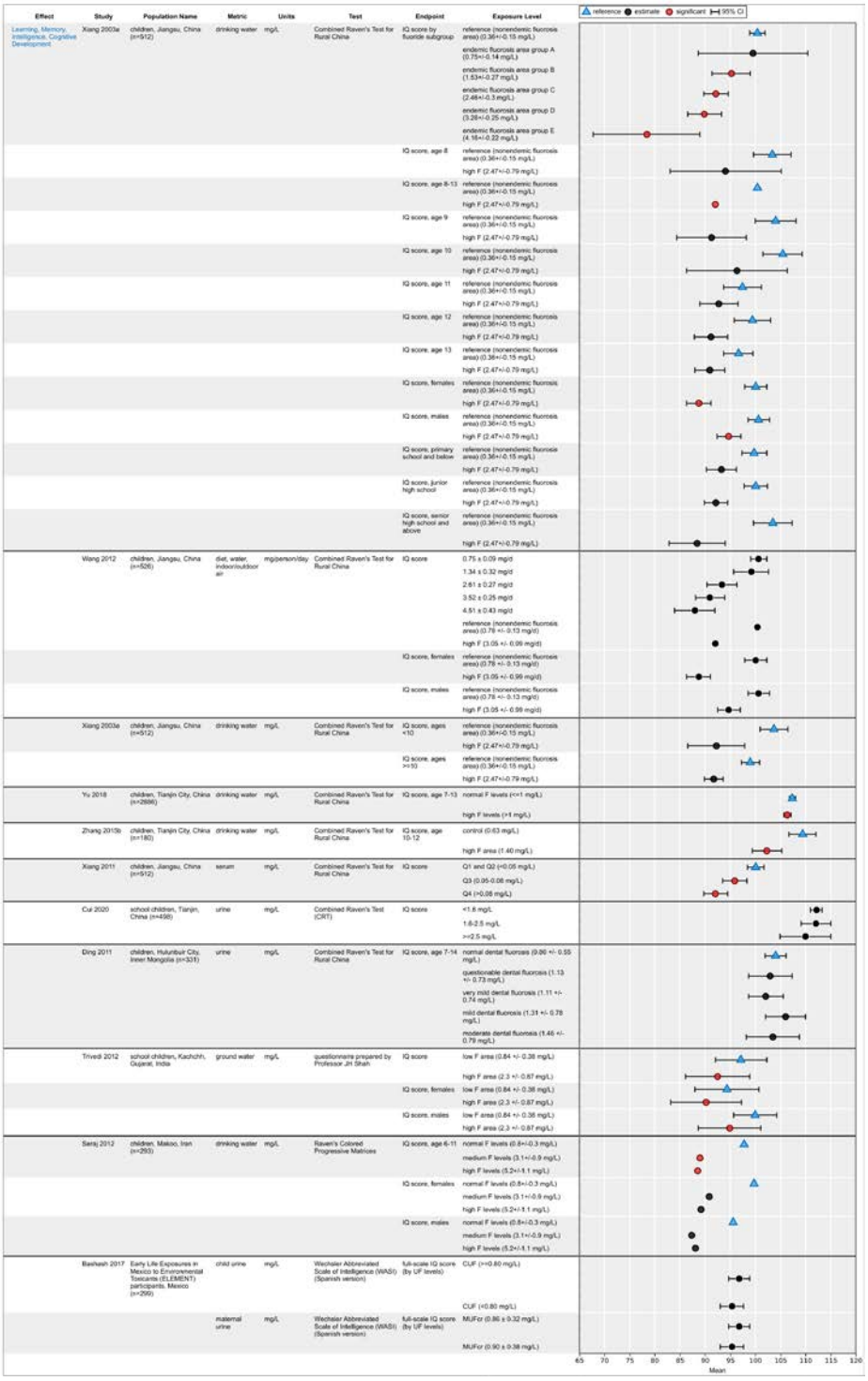
Neurodevelopmental or Cognitive Effects and Outcomes

Figure D1. IQ Distribution in Children by Fluoride Exposure (lower risk-of-bias studies; presented as % in area or % of total group)



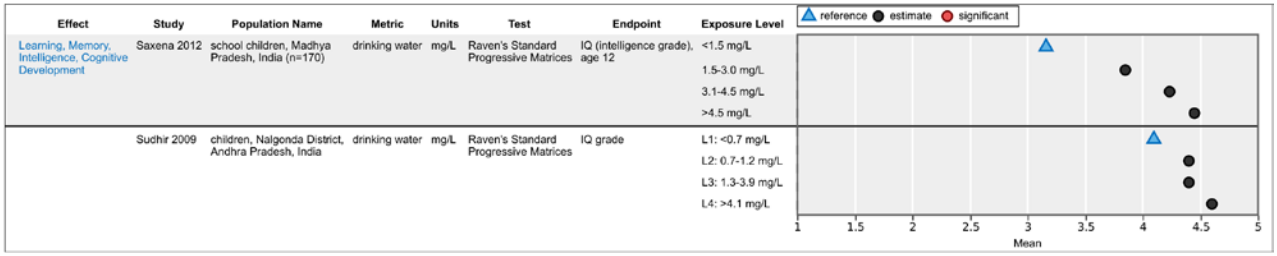
Interactive figure and additional study details in HAWC here. "F" represents fluoride. Differences in intelligence between the reference group and treatment groups were statistically significant although significance was not reported separately for each score level.

Figure D2. Mean IQ in Children by Fluoride Exposure (lower risk-of-bias studies)



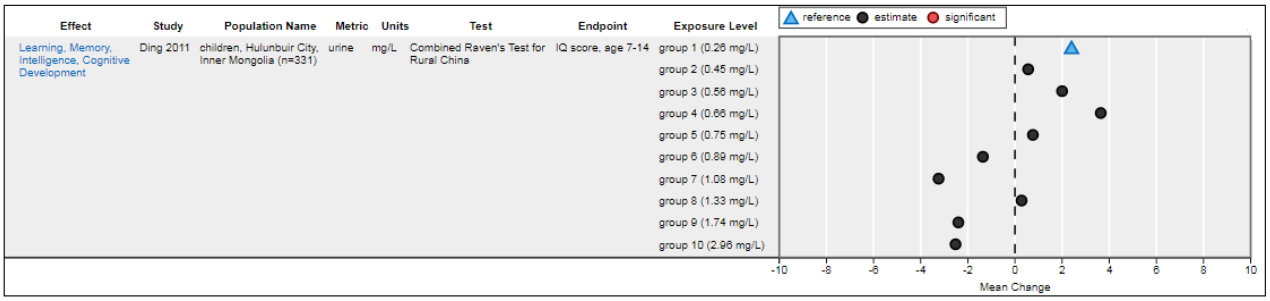
Interactive figure and additional study details in HAWC here. "F" represents fluoride. Three additional publications based on subsample of the larger Yu *et al.* (2018) cohort were identified (Zhao *et al.* 2019, Zhou *et al.* 2019, Zhao *et al.* 2020); however, results from these studies are not presented here. The main study by Yu *et al.* (2018) is considered a better representation of the IQ results. For all studies, SDs are available and can be viewed in HAWC by clicking the data points within the plot area; however, 95% CIs could not be calculated for Seraj *et al.* (2012) because Ns are not available for exposure groups.

Figure D3. Intelligence Grade in Children by Fluoride Exposure (lower risk-of-bias studies; presented as mean)



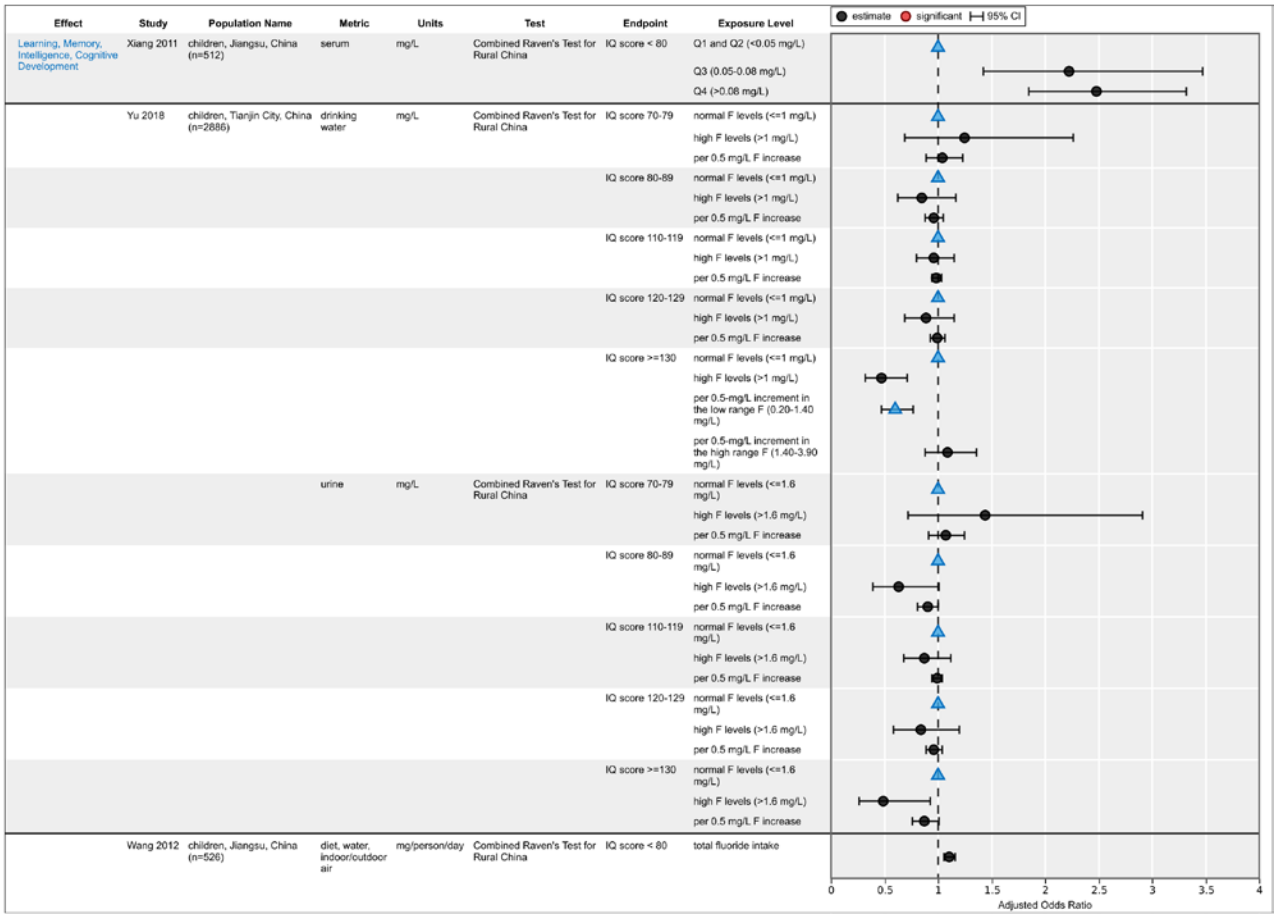
Interactive figure and additional study details in HAWC here. For Saxena *et al.* (2012), children's intelligence was measured using the Raven's Standard Progressive Matrices. Children's scores were converted to percentile and specific grades were allotted based on the percentiles. Grades ranged from intellectually superior (Grade I) to intellectually impaired (Grade V). Results for Soto-Barreras 2019 are not presented here. Outcomes in the study were presented as levels of fluoride exposure associated with each intelligence grade. Results reported were not significant.

Figure D4. Mean Change in IQ in Children by Fluoride Exposure (lower risk-of-bias studies)



Interactive figure and additional study details in HAWC here. For Ding *et al.* (2011), SDs are available and can be viewed in HAWC by clicking the data points within the plot area; however, 95% CIs could not be calculated because Ns for each exposure group are not available.

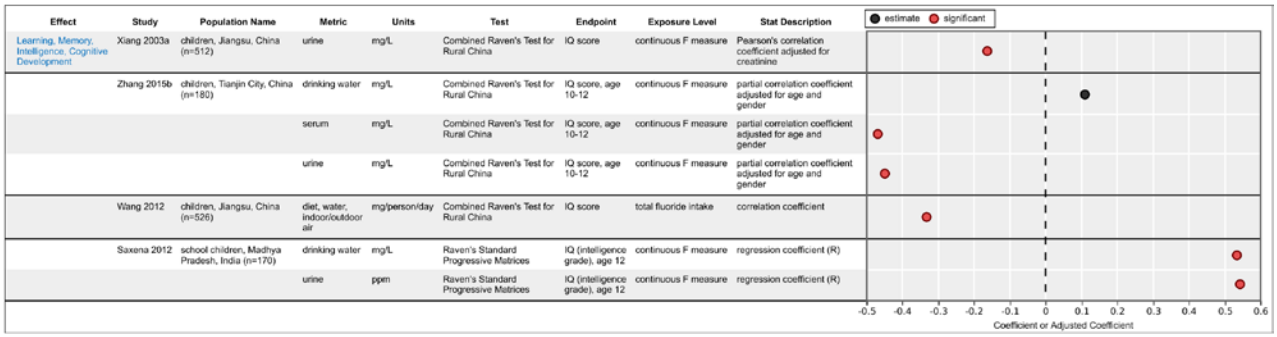
Figure D5. IQ Score in Children by Fluoride Exposure (lower risk-of-bias studies; presented as adjusted OR)



Effect	Study	Population Name	Metric	Units	Test	Endpoint	Exposure Level
Learning, Memory, Intelligence, Cognitive Development	Xiang 2011	children, Jiangsu, China (n=512)	serum	mg/L	Combined Raven's Test for Rural China	IQ score < 80	Q1 and Q2 (<0.05 mg/L)
							Q3 (0.05-0.08 mg/L)
							Q4 (>0.08 mg/L)
	Yu 2018	children, Tianjin City, China (n=2886)	drinking water	mg/L	Combined Raven's Test for Rural China	IQ score 70-79	normal F levels (<=1 mg/L)
							high F levels (>1 mg/L)
							per 0.5 mg/L F increase
						IQ score 80-89	normal F levels (<=1 mg/L)
							high F levels (>1 mg/L)
							per 0.5 mg/L F increase
						IQ score 110-119	normal F levels (<=1 mg/L)
							high F levels (>1 mg/L)
							per 0.5 mg/L F increase
						IQ score 120-129	normal F levels (<=1 mg/L)
							high F levels (>1 mg/L)
							per 0.5 mg/L F increase
						IQ score >=130	normal F levels (<=1 mg/L)
							high F levels (>1 mg/L)
							per 0.5-mg/L increment in the low range F (0.20-1.40 mg/L)
							per 0.5-mg/L increment in the high range F (1.40-3.90 mg/L)
			urine	mg/L	Combined Raven's Test for Rural China	IQ score 70-79	normal F levels (<=1.6 mg/L)
							high F levels (>1.6 mg/L)
							per 0.5 mg/L F increase
						IQ score 80-89	normal F levels (<=1.6 mg/L)
							high F levels (>1.6 mg/L)
							per 0.5 mg/L F increase
						IQ score 110-119	normal F levels (<=1.6 mg/L)
							high F levels (>1.6 mg/L)
							per 0.5 mg/L F increase
						IQ score 120-129	normal F levels (<=1.6 mg/L)
							high F levels (>1.6 mg/L)
							per 0.5 mg/L F increase
						IQ score >=130	normal F levels (<=1.6 mg/L)
							high F levels (>1.6 mg/L)
							per 0.5 mg/L F increase
	Wang 2012	children, Jiangsu, China (n=526)	diet, water, indoor/outdoor air	mg/person/day	Combined Raven's Test for Rural China	IQ score < 80	total fluoride intake

Interactive figure and additional study details in HAWC here. For Xiang *et al.* (2011), there was a significant linear trend across different levels of serum fluoride for IQ score < 80 ($p < 0.001$). For Yu *et al.* (2018), significance levels by IQ score were not reported.

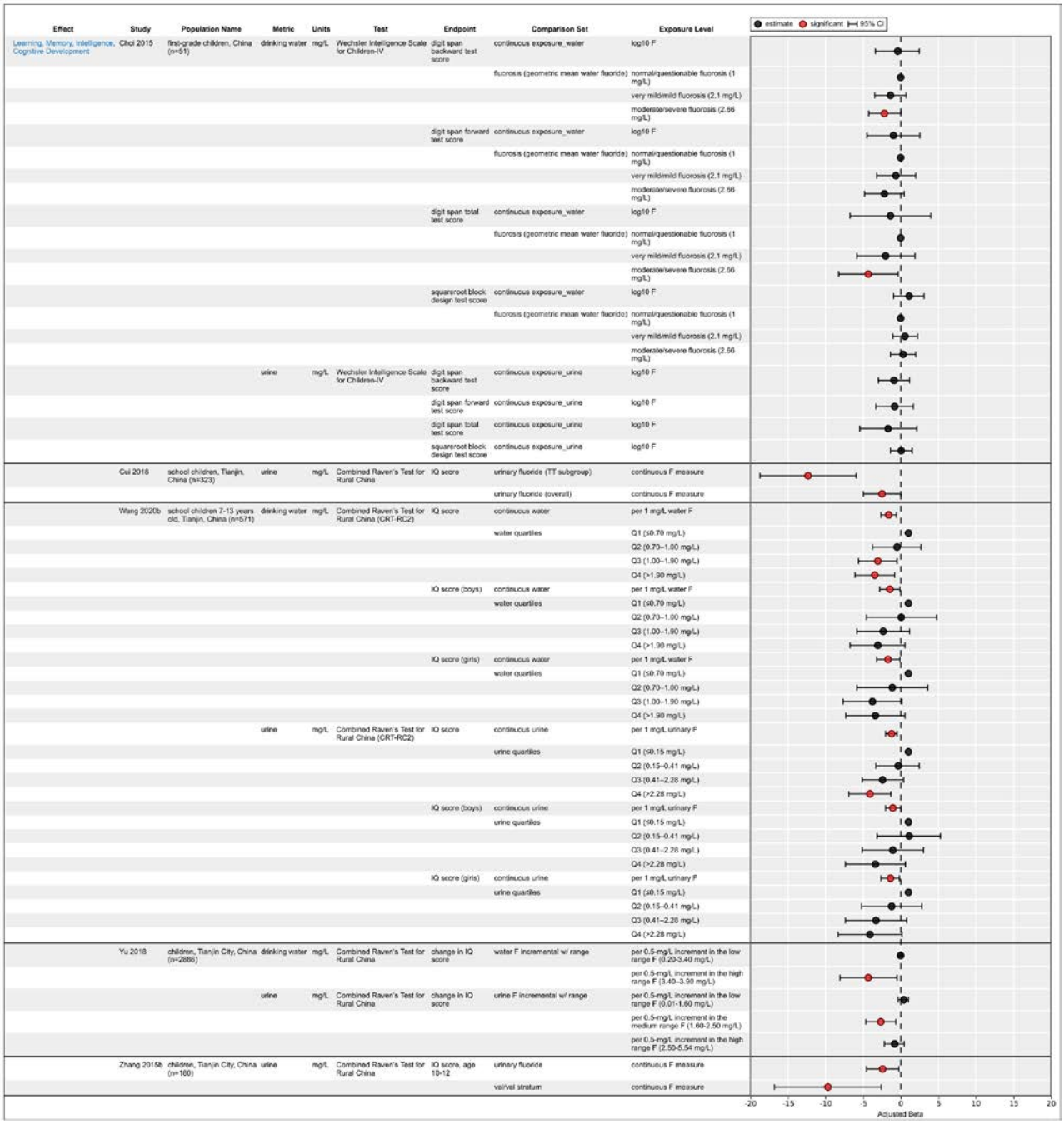
Figure D6. Correlations between IQ Score and Fluoride Exposure in Children (lower risk-of-bias studies; presented as coefficient)



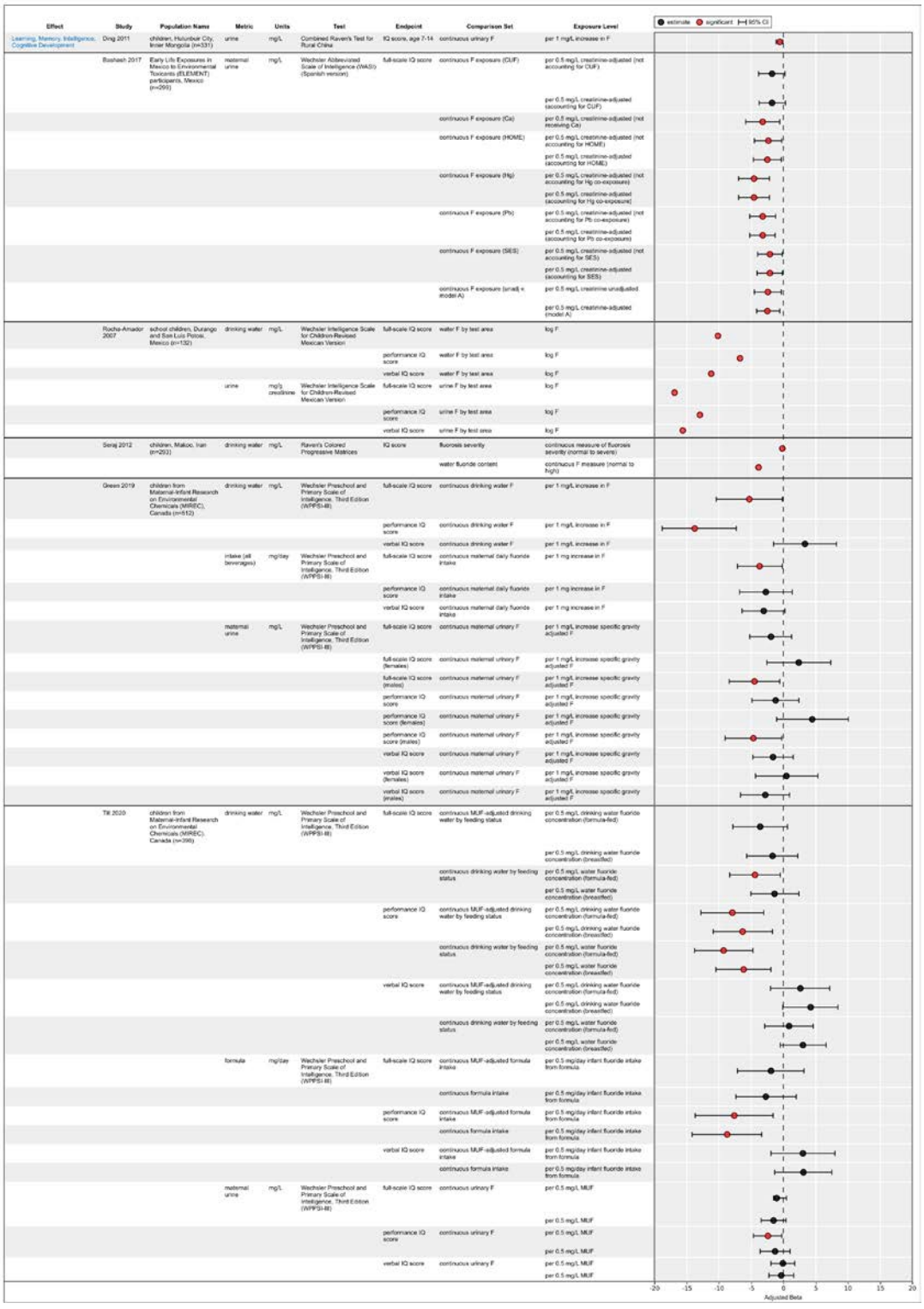
Interactive figure and additional study details in HAWC here. "F" represents fluoride. For Saxena *et al.* (2012), a significant relationship between water fluoride level and intelligence grade was observed. Increasing intelligence grades reflected increasing levels of impairment (reduced intelligence) in children. Zhao *et al.* (2020) and Zhou *et al.* (2019) also had correlations, but these were based on a subsample of the Yu *et al.* (2018) study (which presented betas and provided a better representation of the IQ data).

Figure D7. Correlations between IQ Score and Fluoride Exposure in Children (lower risk-of-bias studies; presented as adjusted beta)—(a) China; (b) all other areas

(a)

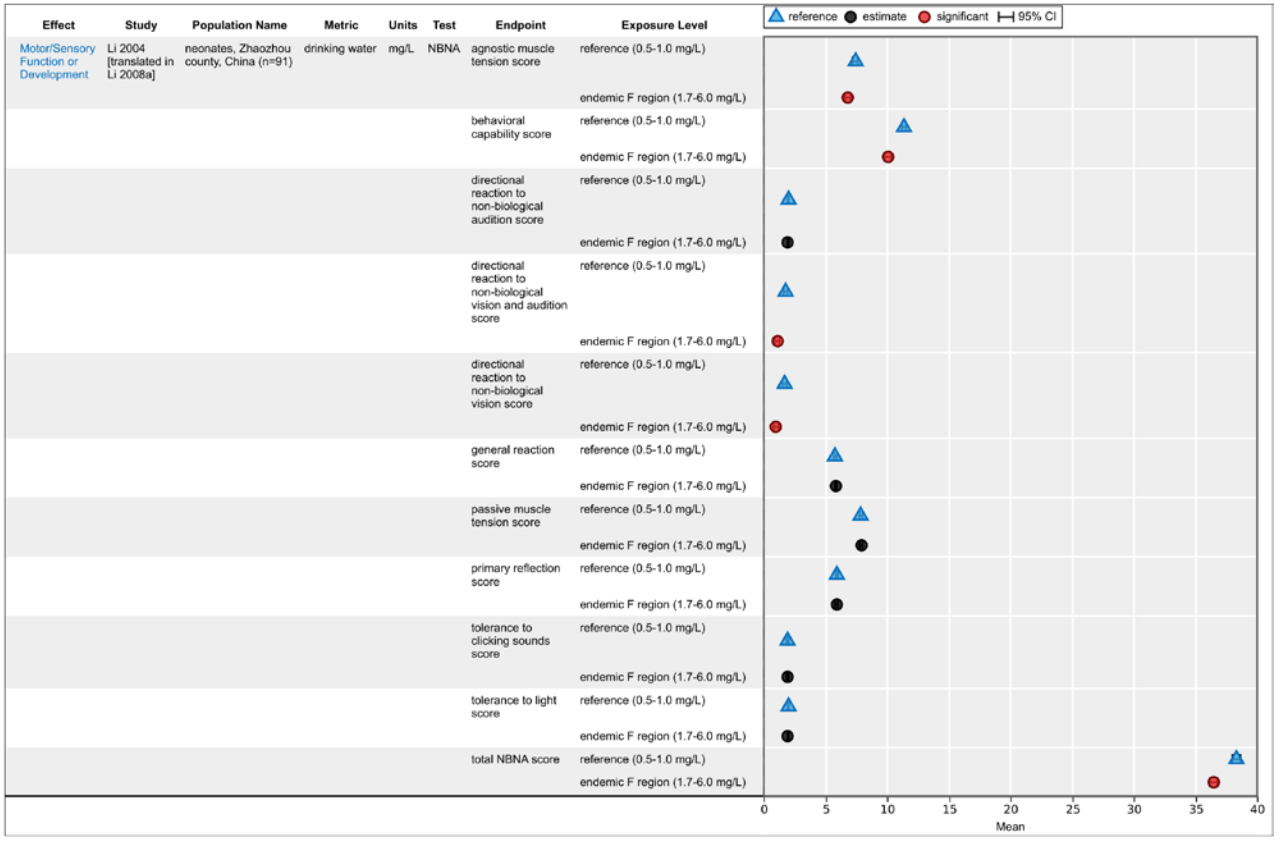


(b)



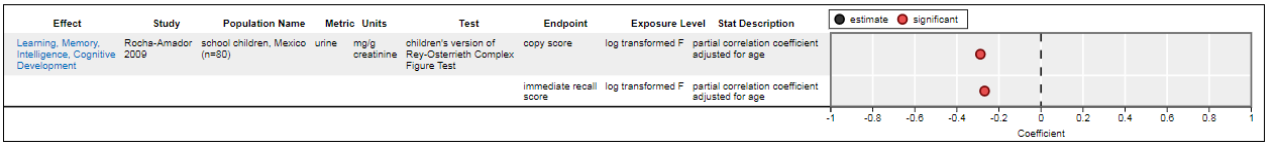
Interactive figure and additional study details in HAWC here for part (a) and here for part (b). "F" represents fluoride. For Yu *et al.* (2018), authors note an obvious decrease in the IQ score at water fluoride exposure levels between 3.40 mg/L and 3.90 mg/L and a similar adverse effect on IQ scores at urinary fluoride exposure levels from 1.60 mg/L to 2.50 mg/L, and so the changes in IQ score are indicated as significant; however, significance levels by change in IQ score were not reported.

Figure D8. Mean Motor/Sensory Scores in Children by Fluoride Exposure (lower risk-of-bias studies)



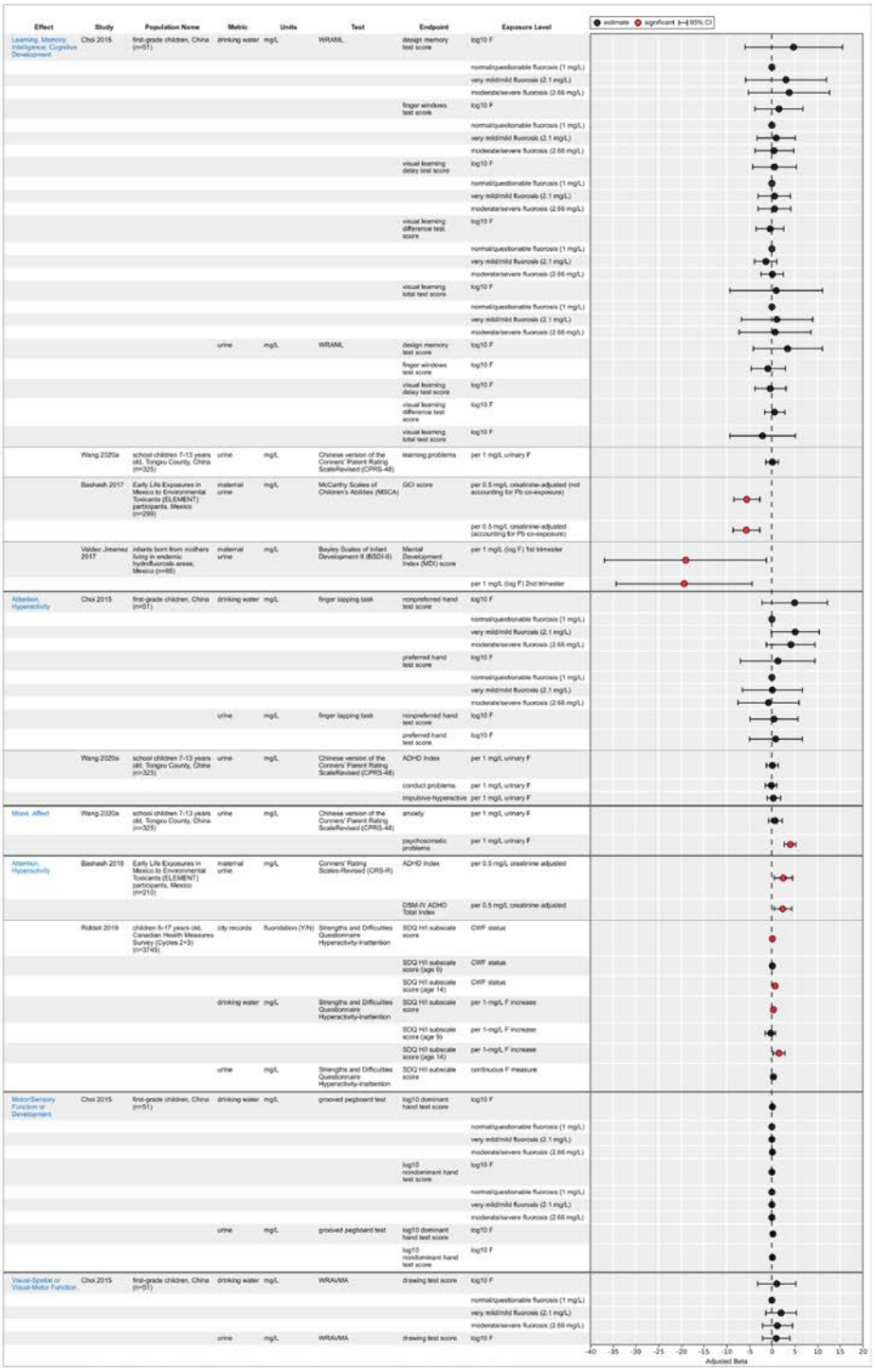
Interactive figure and additional study details in HAWC here. "F" represents fluoride. 95% CIs are small and are within figure symbols and may be difficult to see. Values for SDs and 95% CIs can be viewed in HAWC by clicking the data points within the plot area.

Figure D9. Correlations between Other Neurological Effects and Fluoride Exposure in Children (lower risk-of-bias studies; presented as coefficient)



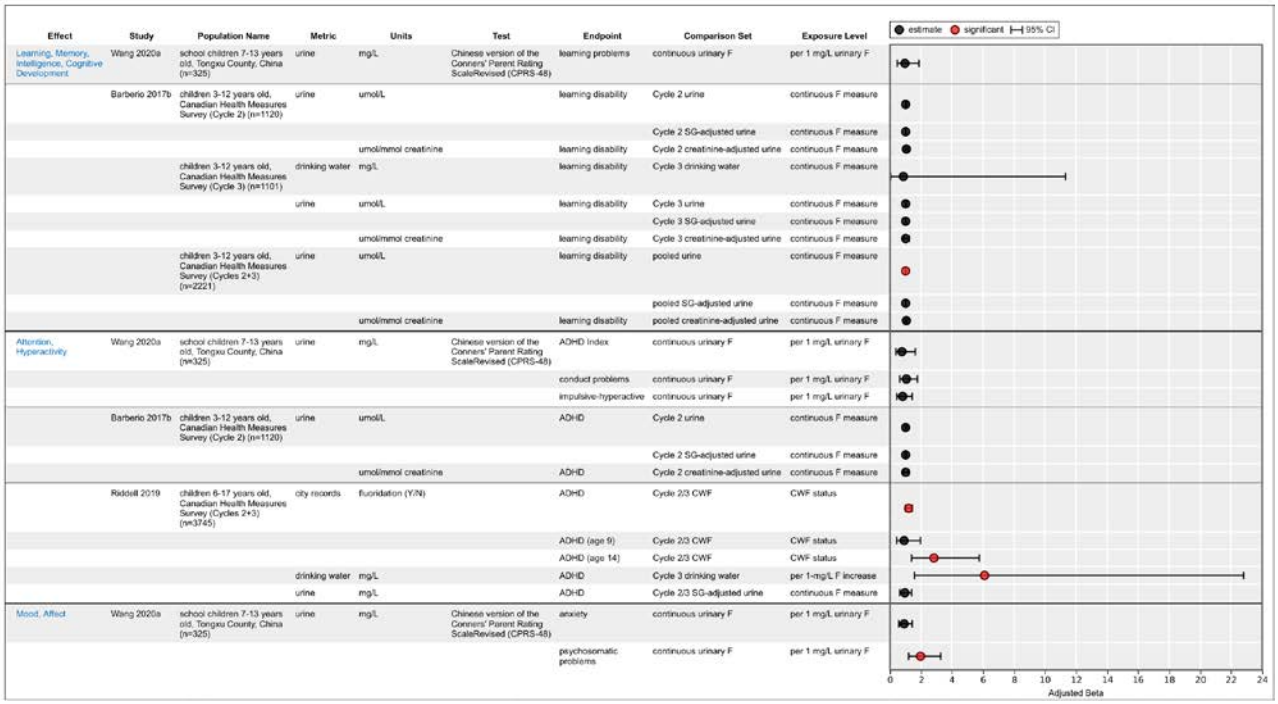
Interactive figure and additional study details in HAWC here. "F" represents fluoride.

Figure D10. Correlations between Other Neurological Effects and Fluoride Exposure in Children (lower risk-of-bias studies; presented as adjusted beta)



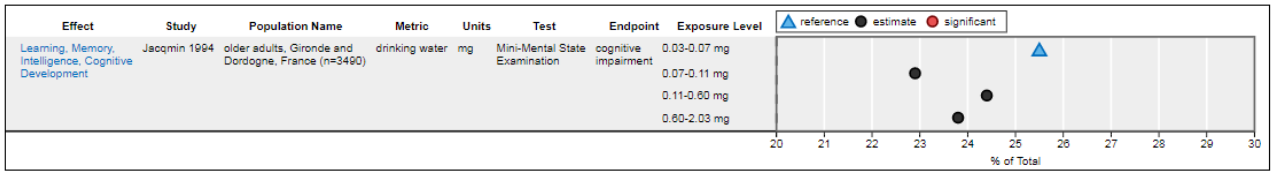
Interactive figure and additional study details in HAWC here. "F" represents fluoride. Bashash *et al.* (2018) observed significant associations between maternal urinary fluoride and ADHD-like symptoms related to inattention (an increase in 0.5 mg/L of maternal urinary fluoride was associated with a 2.84-point increase in the DSM-IV Inattention Index and a 2.54-point increase in Cognitive Problems and Inattention Index). These two scales contributed to the global ADHD Index and the DSM-IV ADHD Total Index shown here.

Figure D11. Correlations between Other Neurological Effects and Fluoride Exposure in Children (lower risk-of-bias studies; presented as adjusted OR)



Interactive figure and additional study details in HAWC here. "F" represents fluoride. Drinking water results for Barberio *et al.* (2017b) have a large confidence interval and are not completely visible in the figure. 95% CIs are 0.068–11.33 and can be viewed in HAWC by clicking the OR within the plot area.

Figure D12. Cognitive Impairment in Adults by Fluoride Exposure (lower risk-of-bias studies; presented as % of total group)



Interactive figure and additional study details in HAWC here. Results from Li *et al.* (2016) suggested that fluoride exposure may be a risk factor for cognitive impairment in elderly subjects; however, results from the study were not conducive to presentation in this visualization.

ABOUT THIS REVIEW

Sources of Support

National Institute of Environmental Health Sciences/Division of the National Toxicology Program

Contributors

Evaluation Team

The evaluation team is composed of federal staff and contractor staff support.

Name	Affiliation
Kyla Taylor, PhD	NIEHS/DNTP, Project Lead
John Bucher, PhD	NIEHS/DNTP
Andrew Rooney, PhD	NIEHS/DNTP
Vickie Walker	NIEHS/DNTP
Cynthia J. Willson, PhD, DVM, DACVP	ILS
Louise Assem, PhD	ICF [f]
Carlye A. Austin, PhD	ICF [d]
Robyn Blain, PhD	ICF [abcdefg]
Natalie Blanton, MPH	ICF [a]
Kristin Bornstein, PhD	ICF [f]
Canden Byrd	ICF [a]
Michelle Cawley	ICF [c]
Jonathan Cohen, PhD	ICF [g]
Ryan Cronk, PhD	ICF [d]
Sorina E. Eftim, PhD	ICF [fg]
Anna Engstrom, PhD	ICF [abcdf]
Jeremy S. Frye, MLS	ICF [a]
Kaitlin Geary	ICF [c]
Susan Goldhaber, MPH	ICF [e]
Ali Goldstone, MPH	ICF [fg]
Pamela Hartman, MEM	ICF [abceg]
Cara Henning, PhD	ICF [a]
Melinda Hoang, MPH	ICF [d]
Tao Hong	ICF [ef]
Jennifer Hsieh, MSPH	ICF [f]
Penelope Kellar	ICF [a]
Courtney Lemeris	ICF [a]
Cynthia J. Lin, PhD MSPH	ICF [g]
Alex Lindahl, MPH	ICF [f]
Kristen Magnuson, MESM	ICF [abcde]
Rachel McGill	ICF [c]
Shannon McGinnis, PhD	ICF [d]

Name	Affiliation
William Mendez, PhD	ICF [b]
Whitney Mitchell	ICF [a]
Revathi Muralidharan	ICF [de]
Johanna Rochester, PhD	ICF [def]
Amanda Ross, MEd	ICF [d]
Pam Ross, MSPH	ICF [df]
Jennifer Seed, PhD	ICF [de]
Karen E. Setty, MESM, PhD	ICF [d]
Codi Sharp	ICF [de]
Robert Shin, MHS	ICF [e]
Kelly Shipkowski, PhD	ICF [df]
Christopher A. Sibrizzi, MPH	ICF [abcdefg]
Raquel A. Silva, PhD	ICF [d]
Samantha J. Snow, PhD, DABT	ICF [ef]
Parnian Soleymani, MS	ICF [d]
Anna Stamatogiannakis	ICF [a]
Nicole Vetter, MLS	ICF [a]
River Williams, BS	ICF [c]

Note: the roles of individual contractors differed: [a] indicates monograph development; [b] indicates review of data, results, and analyses; [c] indicates database and HAWC support; [d] indicates literature screening, [e] indicates data extraction, [f] indicates risk-of-bias assessment, [g] indicates meta-analysis

Peer Reviewers

The peer reviewers were outside experts selected for their experience with fluoride, developmental neurobehavioral toxicity, and systematic review procedures. Peer reviewers were screened for conflict of interest prior to their service and did not report any conflicts of interest. Service as a peer reviewer does not necessarily indicate that the reviewer endorses the final document.

Protocol Reviewers

Name	Affiliation
Joseph Braun, PhD [a]	Brown University
Marie Sutton, PhD [a]	Health Research Board
Thomas Zoeller, PhD [a]	University of Massachusetts, Amherst
Thomas Webster, PhD [a]	Boston University
Gail Wasserman, PhD [a]	Columbia University
Suril Mehta, MPH [b]	NIEHS
Tianjing Li, PhD [b]	University of Colorado Denver

[a] Reviewer of the systematic review protocol; [b] Reviewer of meta-analysis protocol
no conflicts of interest declared

Technical Review of Draft Monograph

Name	Affiliation
Freya Kamal, PhD	NIEHS (retired)

no conflicts of interest declared

Protocol History and Revisions

Date	Activity or revision
December 14, 2016	**Draft evaluation protocol reviewed:** sent to technical advisors for peer review
April 10, 2017	**Draft human risk-of-bias protocol reviewed;** sent to technical advisors for peer review
May 2, 2017	**Draft animal risk-of-bias protocol reviewed;** sent to technical advisors for peer review
June 2017	**Evaluation protocol finalized:** Review protocol finalized for use and posting
May 2019	**Revised protocol:** Revised review protocol posted
September 2020	**Revised protocol:** Revised review protocol posted

APPENDICES

Appendix 1. Literature Search Strategy

The strategy for this search is broad for the consideration of neurodevelopmental or cognitive endpoints and comprehensive for fluoride as an exposure or treatment in order to ensure inclusion of relevant papers. The search terms for PubMed are provided below. The specific search strategies for other databases are available in the protocol (https://ntp.niehs.nih.gov/go/785076).

Database	Search Terms
PUBMED	((Fluorides[mh:noexp] OR fluorides, topical[mh] OR sodium fluoride[mh] OR Fluorosis, Dental[mh] OR fluorosis[tiab] OR fluorid*[tiab] OR flurid*[tiab] OR fluorin*[tiab] OR florin*[tiab]) NOT (18F[tiab] OR f-18[tiab] OR 19F[tiab] OR f-19[tiab] OR f-labeled[tiab] OR "fluorine-18"[tiab] OR "fluorine-19"[tiab] OR pet-scan[tiab] OR radioligand*[tiab])) AND ((Aryl Hydrocarbon Hydroxylases[mh] OR Aryl Hydrocarbon Receptor Nuclear Translocator[mh] OR Behavior and Behavior Mechanisms[mh] OR Gene Expression Regulation[mh] OR Glucuronosyltransferase[mh] OR Intelligence tests[mh] OR Malate Dehydrogenase[mh] OR Mediator Complex Subunit 1[mh] OR Mental disorders[mh] OR Mental processes[mh] OR Monocarboxylic Acid Transporters[mh] OR Myelin Basic Protein[mh] OR nervous system[mh] OR nervous system diseases[mh] OR nervous system physiological phenomena[mh] OR Neurogranin[mh] OR Oligodendroglia[mh] OR Peroxisome Proliferator-Activated Receptors[mh] OR Psychological Phenomena and Processes[mh] OR Receptors, thyroid hormone[mh] OR Receptors, thyrotropin[mh] OR Retinoid X Receptors[mh] OR thyroid diseases[mh] OR thyroid hormones[mh] OR Thyrotropin-releasing hormone[mh] OR Thyroxine-Binding Proteins[mh] OR Pregnane X Receptor[supplementary concept] OR thyroid-hormone-receptor interacting protein[supplementary concept] OR Constitutive androstane receptor[supplementary concept] OR Academic performance[tiab] OR auditory[tiab] OR cortical[tiab] OR delayed development[tiab] OR developmental impairment[tiab] OR developmental-delay*[tiab] OR developmental-disorder*[tiab] OR euthyroid[tiab] OR gait[tiab] OR glia*[tiab] OR gliogenesis[tiab] OR hyperactiv*[tiab] OR impulse-control[tiab] OR iodide peroxidase[tiab] OR IQ[tiab] OR ischemi*[tiab] OR locomotor[tiab] OR mental deficiency[tiab] OR mental development[tiab] OR mental illness[tiab] OR mental-deficit[tiab] OR mobility[tiab] OR mood[tiab] OR morris-maze[tiab] OR morris-water[tiab] OR motor abilit*[tiab] OR Motor activities[tiab] OR motor performance[tiab] OR nerve[tiab] OR neural[tiab] OR neurobehav*[tiab] OR Neurocognitive impairment[tiab] OR neurodegenerat*[tiab] OR Neurodevelopment*[tiab] OR neurodisease*[tiab] OR neurologic*[tiab] OR neuromuscular[tiab] OR neuron*[tiab] OR neuropath*[tiab] OR obsessive compulsive[tiab] OR OCD[tiab] OR olfaction[tiab] OR olfactory[tiab] OR open-field-test[tiab] OR passive avoidance[tiab] OR plasticity[tiab] OR senil*[tiab] OR sociab*[tiab] OR speech*[tiab] OR spelling[tiab] OR stereotypic-movement*[tiab] OR synap*[tiab] OR tauopath*[tiab] OR Thyroglobulin[tiab] OR Thyroid disease*[tiab] OR thyroid gland[tiab] OR thyroid hormone*[tiab] OR thyronine*[tiab] OR visual motor[tiab] OR Visuospatial processing[tiab] OR water maze[tiab]) OR ((active-avoidance[tiab] OR ADHD[tiab] OR alzheimer*[tiab] OR amygdala[tiab] OR antisocial[tiab] OR anxiety[tiab] OR anxious[tiab] OR asperger*[tiab] OR attention deficit[tiab] OR autism[tiab] OR autistic[tiab] OR behavioral[tiab] OR behaviors[tiab] OR behavioural[tiab] OR behaviours[tiab] OR bipolar[tiab] OR cerebellum[tiab] OR cognition[tiab] OR cognitive[tiab] OR communication-disorder*[tiab] OR comprehension[tiab] OR cranial[tiab] OR dementia[tiab] OR dendrit*[tiab] OR dentate-gyrus[tiab] OR depression[tiab] OR dextrothyroxine[tiab] OR diiodothyronine*[tiab] OR diiodotyrosine[tiab] OR down syndrome[tiab] OR dyslexia[tiab] OR entorhinal cortex[tiab] OR epilep*[tiab] OR gangli*[tiab] OR goiter[tiab] OR graves-disease[tiab] OR hearing[tiab] OR hippocamp*[tiab] OR human development[tiab] OR hyperthyroid*[tiab] OR hypothalam*[tiab] OR hypothyroid*[tiab] OR impulsiv*[tiab] OR Intellectual

Database	Search Terms
	disability[tiab] OR intelligence[tiab] OR language[tiab] OR learning[tiab] OR lewy bod*[tiab] OR long-term potentiation[tiab] OR long-term synaptic depression[tiab] OR memory[tiab] OR mental disorder*[tiab] OR mental recall[tiab] OR monoiodotyrosine[tiab] OR Motor activity[tiab] OR motor skill*[tiab] OR multiple sclerosis[tiab] OR myxedema[tiab] OR Nervous system[tiab] OR nervous-system[tiab] OR neurit*[tiab] OR optic[tiab] OR palsy[tiab] OR panic[tiab] OR parahippocamp*[tiab] OR paranoia[tiab] OR paranoid[tiab] OR parkinson*[tiab] OR perception[tiab] OR perforant*[tiab] OR personality[tiab] OR phobia[tiab] OR problem solving[tiab] OR proprioception[tiab] OR psychomotor[tiab] OR reflex[tiab] OR risk taking[tiab] OR schizophrenia[tiab] OR seizure*[tiab] OR sensation*[tiab] OR sleep[tiab] OR smell[tiab] OR spatial behavior[tiab] OR stroke[tiab] OR substantia-nigra[tiab] OR taste[tiab] OR thyroiditis[tiab] OR thyrotoxicosis[tiab] OR Thyrotropin[tiab] OR thyroxine[tiab] OR triiodothyronine[tiab] OR vision[tiab]) NOT medline[sb]))

Appendix 2. List of Included Studies

Studies in Humans

As described in **Figure 4**, 159 human studies were included; however, full data extraction was only conducted on studies with neurological outcomes or thyroid hormone data. Data extraction was completed using HAWC. Data were extracted from a subset of included studies in humans (n = 116) and are available in HAWC based on outcome. The following lists of references are organized as studies that are available in HAWC followed by studies that are not available in HAWC. Specifically, data for primary neurodevelopmental or cognitive outcomes (learning, memory, and intelligence) and secondary neurobehavioral outcomes (anxiety, aggression, motor activity, or biochemical changes), as well as thyroid hormone level data, were extracted from included human studies and are available in HAWC. Data for included studies identified through the 2020 literature search update were only extracted for primary neurodevelopmental or cognitive outcomes; a subset of these studies (n = 5) also included secondary neurobehavioral outcomes and/or thyroid hormone level data that were not extracted because those data would not materially advance the human or mechanistic findings. Included human studies that only evaluated other thyroid-related effects such as goiters or thyroid size (n = 43) were not extracted and are not available in HAWC. The list below presents the 159 human studies that were included in the review. An overview of the screening results is outlined in the study selection diagram (**Figure 4**) that reports numbers of included studies as well as numbers of studies excluded for each reason at the full text review stage.

Studies Available in HAWC

An J, Mei S, Liu A, Fu Y, Wang C. 1992. [Effect of high level of fluoride on children's intelligence]. *Chin J Control Endem Dis* 7(2): 93-94.

Aravind A, Dhanya RS, Narayan A, Sam G, Adarsh VJ, Kiran M. 2016. Effect of fluoridated water on intelligence in 10-12-year-old school children. *J Int Soc Prev Community Dent* 6(Suppl 3): S237-S242.

Bai A, Li Y, Fan Z, Li X, Li P. 2014. [Intelligence and growth development of children in coal-burning-borne arsenism and fluorosis areas: An investigation study]. *Chin J Endemiol* 33(2): 160-163.

Barberio AM, Hosein FS, Quinonez C, McLaren L. 2017. Fluoride exposure and indicators of thyroid functioning in the Canadian population: Implications for community water fluoridation. *J Epidemiol Community Health* 71: 1019-1025.

Barberio AM, Quinonez C, Hosein FS, McLaren L. 2017. Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation. *Can J Public Health* 108: 229-239.

Bashash M, Thomas D, Hu H, Martinez-Mier EA, Sanchez BN, Basu N, Peterson KE, Ettinger AS, Wright R, Zhang Z, Liu Y, Schnaas L, Mercado-Garcia A, Tellez-Rojo MM, Hernandez-Avila M. 2017. Prenatal fluoride exposure and cognitive outcomes in children at 4 and 6-12 years of age in Mexico. *Environ Health Perspect* 125(9): 1-12.

Bashash M, Marchand M, Hu H, Till C, Martinez-Mier EA, Sanchez BN, Basu N, Peterson KE, Green R, Schnaas L, Mercado-Garcia A, Hernandez-Avila M, Tellez-Rojo MM. 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.

Broadbent JM, Thomson WM, Moffitt TE, Poulton R. 2015. Community water fluoridation and intelligence response. *Am J Public Health* 105: 3-4.

Chen YX, Han FL, Zhoua ZL, Zhang HQ, Jiao XS, Zhang SC, Huang MC, Chang TQ, Dong YF. 1991. [Research on the intellectual development of children in high fluoride areas]. *Chin J Control Endem Dis* 6(Suppl): 99-100.

Chen YX, Han FL, Zhoua ZL, Zhang HQ, Jiao XS, Zhang SC, Huang MC, Chang TQ, Dong YF. 2008. Research on the intellectual development of children in high fluoride areas. *Fluoride* 41: 120-124.

Chinoy NJ, Narayana MV. 1992. Studies on fluorosis in Mehsana District of North Gujarat. *Proc Zool Soc* 45: 157-161.

Choi AL, Zhang Y, Sun G, Bellinger DC, Wang K, Yang XJ, Li JS, Zheng Q, Fu Y, Grandjean P. 2015. Association of lifetime exposure to fluoride and cognitive functions in Chinese children: A pilot study. *Neurotoxicol Teratol* 47: 96-101.

Cui Y, Zhang B, Ma J, Wang Y, Zhao L, Hou C, Yu J, Zhao Y, Zhang Z, Nie J, Gao T, Zhou G, Liu H. 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.

Cui Y, Yu J, Zhang B, Guo B, Gao T, Liu H. 2020. The relationships between thyroid-stimulating hormone and/or dopamine levels in peripheral blood and IQ in children with different urinary iodine concentrations. *Neurosci Lett* 729: 134981.

Das K, Mondal NK. 2016. Dental fluorosis and urinary fluoride concentration as a reflection of fluoride exposure and its impact on IQ level and BMI of children of Laxmisagar, Simlapal Block of Bankura District, W.B., India. *Environ Monit Assess* 188: 218.

Ding Y, Sun H, Han H, Wang W, Ji X, Liu X, Sun D. 2011. The relationships between low levels of urine fluoride on children's intelligence, dental fluorosis in endemic fluorosis areas in Hulunbuir, Inner Mongolia, China. *J Hazard Mater* 186: 1942-1946.

Du L, Wan C, Cao X, Liu J. 1992. [The effect of fluorine on the developing human brain]. *Chin J Pathol* 21(4): 218-220.

Du L, Wan C, Cao X, Liu J. 2008. The effect of fluorine on the developing human brain. *Fluoride* 41: 327-330.

Duan J, Zhao M, Wang L, Fang D, Wang Y, Wang W. 1995. A comparative analysis of the results of multiple tests in patients with chronic industrial fluorosis. Guizhou Med J 18(3): 179-180.
Erickson JD, Hay S. 1976. Water fluoridation and congenital malformations: No association. *J Am Dent Assoc* 93: 981-984.

Erickson JD, Hay S. 1976. Water fluoridation and congenital malformations: No association. *J Am Dent Assoc* 93: 981-984.

Erickson JD. 1980. Down syndrome, water fluoridation, and maternal age. *Teratology* 21: 177-180.

Eswar P, Nagesh L, Devaraj CG. 2011. Intelligent quotients of 12-14 year old school children in a high and low fluoride village in India. *Fluoride* 44: 168-172.

Fan Z, Dai H, Bai A, Li P, Li T, Li G. 2007. Effect of high fluoride exposure in children's intelligence. *J Environ Health* 24(10): 802-803.

Green R, Lanphear B, Hornung R, Flora D, Martinez-Mier EA, Neufeld R, Ayotte P, Muckle G, Till C. 2019. Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada. *JAMA Pediatr*: E1-E9.

Guo XC, Wang RY, Cheng CF, Wei WS, Tang LM, Wang QS, Tang DX, Liu GW, He GD, Li SL. 1991. [A preliminary investigation of the IQs of 7-13 year-old children from an area with coal burning-related fluoride poisoning]. *Chin J Epidemiol* 10(2): 98-100.

Guo XC, Wang RY, Cheng CF, Wei WS, Tang LM, Wang QS, Tang DX, Liu GW, He GD, Li SL. 2008. A preliminary investigation of the IQs of 7-13 year-old children from an area with coal burning-related fluoride poisoning. *Fluoride* 41: 125-128.

Guo ZY, He YH, Zhu QX. 2001. [Research on the neurobehavioral function of workers occupationally exposed to fluoride]. *Ind Hlth & Occup Dis* 27(6): 346-348.

Guo ZY, He YH, Zhu QX. 2008. Research on the neurobehavioral function of workers occupationally exposed to fluoride. *Fluoride* 41: 152-155.

He H, Cheng ZS, Liu WQ. 1989. [Effects of fluorine on the human fetus]. *J Control Endem Dis* 4(3): 136-138.

He H, Cheng ZS, Liu WQ. 2008. Effects of fluorine on the human fetus. *Fluoride* 41: 321-326.

He MX, Zhang CN. 2010. [Investigation of children's intelligence quotient and dental fluorosis in drinking water-type of endemic fluorosis area in Pucheng County, Shaanxi Province before and after drinking water change]. *Chin J Endemiol* 29: 547-548.

Hong FG, Cao YX, Yang D, Wang H. 2001. [Research on the effects of fluoride on child intellectual development under different environmental conditions]. *Chin Prim Health Care* 15(3): 56-57.

Hong FG, Cao YX, Yang D, Wang H. 2008. Research on the effects of fluoride on child intellectual development under different environmental conditions. *Fluoride* 41: 156-160.

Hosur MB, Puranik RS, Vanaki S, Puranik SR. 2012. Study of thyroid hormones free triiodothyronine (FT3), free thyroxine (FT4) and thyroid stimulating hormone (TSH) in subjects with dental fluorosis. *Eur J Dent* 6: 184-190.

Jacqmin H, Commenges D, Letenneur L, Barberger-Gateau P, Dartigues JF. 1994. Components of drinking water and risk of cognitive impairment in the elderly. *Am J Epidemiol* 139: 48-57.

Kang J, Cheng Y, Wu K, Lin S, He G, Jin Y. 2011. Effect of exposure to fluoride and arsenic in drinking water of Hangjinhouqi on children's intelligence. *Chinese School Health*: 679-681.

Karimzade S, Aghaei M, Mahvi AH. 2014. Investigation of intelligence quotient in 9-12 year-old children exposed to high- and low-drinking water fluoride in West Azerbaijan Province, Iran. *Fluoride* 47: 9-14.

Khan SA, Singh RK, Navit S, Chadha D, Johri N, Navit P, Sharma A, Bahuguna R. 2015. Relationship between dental fluorosis and intelligence quotient of school going children in and around Lucknow District: A cross-sectional study. *J Clin Diagn Res* 9(11): 10-15.

Khandare AL, Gourineni SR, Validandi V. 2017. Dental fluorosis, nutritional status, kidney damage, and thyroid function along with bone metabolic indicators in school-going children living in fluoride-affected hilly areas of Doda District, Jammu and Kashmir, India. *Environ Monit Assess* 189: 579.

Khandare AL, Validandi V, Gourineni SR, Gopalan V, Nagalla B. 2018. Dose-dependent effect of fluoride on clinical and subclinical indices of fluorosis in school going children and its mitigation by supply of safe drinking water for 5 years: An Indian study. *Environ Monit Assess* 190: 110.

Kheradpisheh Z, Mahvi AH, Mirzaei M, Mokhtari M, Azizi R, Fallahzadeh H, Ehrampoush MH. 2018. Correlation between drinking water fluoride and TSH hormone by ANNs and ANFIS. *J Environ Health Sci Eng* 16(1): 11-18.

Kheradpisheh Z, Mirzaei M, Mahvi AH, Mokhtari M, Azizi R, Fallahzadeh H, Ehrampoush MH. 2018. Impact of drinking water fluoride on human thyroid hormones: A case-control study. *Sci Rep* 8: 2674.

Kumar V, Chahar P, Kajjari S, Rahman F, Bansal DK, Kapadia JM. 2018. Fluoride, thyroid hormone derangements and its correlation with tooth eruption pattern among the pediatric population from endemic and non-endemic fluorosis areas. *J Contemp Dent Pract* 19(12): 1512-1516.

Kundu H, Basavaraj P, Singla A, Gupta R, Singh K, Jain S. 2015. Effect of fluoride in drinking water on children's intelligence in high and low fluoride areas of Delhi. *J Indian Assoc Public Health Dent* 13(2): 116-121.

Lamberg M, Hausen H, Vartiainen T. 1997. Symptoms experienced during periods of actual and supposed water fluoridation. *Community Dent Oral Epidemiol* 25: 291-295.

Li F, Chen X, Huang R, Xie Y. 2009. The impact of endemic fluorosis caused by the burning of coal on the development of intelligence in children. *J Environ Health* 26(4): 838-840.

Li J, Yao L, Shao QL, Wu CY. 2004. [Effects of high fluoride level on neonatal neurobehavioral development]. *Chin J Endemiol* 23(5): 463-465.

Li J, Yao L, Shao QL, Wu CY. 2008. Effects of high fluoride level on neonatal neurobehavioral development. *Fluoride* 41: 165-170.

Li M, Gao Y, Cui J, Li Y, Li B, Liu Y, Sun J, Liu X, Liu H, Zhao L, Sun D. 2016. Cognitive impairment and risk factors in elderly people living in fluorosis areas in China. *Biol Trace Elem Res* 172: 53-60.

Li X, Hou G, Yu B, Yuan C, Liu Y, L Z. 2010. Investigation and analysis of children's intelligence and dental fluorosis in high fluoride area. *J Med Pest Control* 26(3): 230-231.

Li XS, Zhi JL, Gao RO. 1995. Effect of fluoride exposure on the intelligence of children. *Fluoride* 28: 189-192.

Li Y, Li X, Wei S. 1994. [Effects of high fluoride intake on child mental work capacity: Preliminary investigation into the mechanisms involved]. *J West Chin University Med Sci* 25(2): 188-191.

Li Y, Li X, Wei S. 2008. Effects of high fluoride intake on child mental work capacity: Preliminary investigation into the mechanisms involved. *Fluoride* 41: 331-335.

Li YP, Jing XY, Chen D, Lin L, Wang ZJ. 2003. [Effects of endemic fluoride poisoning on the intellectual development of children in Baotou]. *Chin J Public Health Manag* 19(4): 337-338.

Li YP, Jing XY, Chen D, Lin L, Wang ZJ. 2008. Effects of endemic fluoride poisoning on the intellectual development of children in Baotou. *Fluoride* 41: 161-164.

Lin F, Ai H, Zhao H, Lin J, Jhiang J, Maimaiti. 1991. High fluoride and low iodine environment and subclinical cretinism in Xinjiang. *Endem Dis Bull* 6(2): 62-67.

Liu S, Lu Y, Sun Z, Wu L, Wang X, Yan S. 2000. [Report on the intellectual ability of children living in high-fluoride water areas]. *Chin J Control Endem Dis* 15(4): 231-232.

Liu SL, Lu Y, Sun ZR, Wu L, Lu WL, Wang XW, Yan S. 2008. Report on the intellectual ability of children living in high-fluoride water areas. *Fluoride* 41: 144-147.

Lu F, Zhang Y, Trivedi A, Jiang X, Chandra D, Zheng J, Nakano Y, Uyghurturk DA, Jalai R, Guner S, Mentes A, DenBesten PK. 2019. Fluoride related changes in behavioral outcomes may relate to increased serotonin. *Physiol Behav* 206: 76-83.

Lu Y, Sun ZR, Wu LN, Wang X, Lu W, Liu SS. 2000. Effect of high-fluoride water on intelligence in children. *Fluoride* 33: 74-78.

Malin AJ, 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.

Malin AJ, 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.

Michael M, Barot VV, Chinoy NJ. 1996. Investigations of soft tissue functions in fluorotic individuals of North Gujarat. *Fluoride* 29: 63-71.

Mondal D, Dutta G, Gupta S. 2016. Inferring the fluoride hydrogeochemistry and effect of consuming fluoride-contaminated drinking water on human health in some endemic areas of Birbhum District, West Bengal. *Environ Geochem Health* 38: 557-576.

Morgan L, Allred E, Tavares M, Bellinger D, Needleman H. 1998. Investigation of the possible associations between fluorosis, fluoride exposure, and childhood behavior problems. *Pediatr Dent* 20: 244-252.

Mustafa DE, Younis UM, Elhag SAA. 2018. The relationship between the fluoride levels in drinking water and the schooling performance of children in rural areas of Khartoum State, Sudan. *Fluoride* 51(2): 102-113.

Nagarajappa R, Pujara P, Sharda AJ, Asawa K, Tak M, Aapaliya P, Bhanushali N. 2013. Comparative assessment of intelligence quotient among children living in high and low fluoride areas of Kutch, India: A pilot study. *Iran J Public Health* 42: 813-818.

Peckham S, Lowery D, Spencer S. 2015. Are fluoride levels in drinking water associated with hypothyroidism prevalence in England? A large observational study of GP practice data and fluoride levels in drinking water. *J Epidemiol Community Health* 69: 619-624.

Poureslami HR, Horri A, Garrusi B. 2011. A comparative study of the IQ of children age 7-9 in a high and a low fluoride water city in Iran. *Fluoride* 44: 163-167.

Qin LS, Huo SY, Chen RL, Chang YZ, Zhao MY. 1990. [Using the Raven's Standard Progressive Matrices to determine the effects of the level of fluoride in drinking water on the intellectual ability of school-age children]. *J Control Endem Dis* 5(4): 203-204.

Qin LS, Huo SY, Chen RL, Chang YZ, Zhao MY. 2008. Using the Raven's Standard Progressive Matrices to determine the effects of the level of fluoride in drinking water on the intellectual ability of school-age children. *Fluoride* 41: 115-119.

Ratan SK, Rattan KN, Pandey RM, Singhal S, Kharab S, Bala M, Singh V, Jhanwar A. 2008. Evaluation of the levels of folate, vitamin B12, homocysteine and fluoride in the parents and the affected neonates with neural tube defect and their matched controls. *Pediatr Surg Int* 24: 803-808.

Razdan P, Patthi B, Kumar JK, Agnihotri N, Chaudhari P, Prasad M. 2017. Effect of fluoride concentration in drinking water on intelligence quotient of 12-14-year-old children in Mathura District: A cross-sectional study. *J Int Soc Prev Community Dent* 7: 252-258.

Ren D, Li K, Liu D. 1989. [A study of the intellectual ability of 8-14 year-old children in high fluoride, low iodine areas]. *Chin J Control Endem Dis* 4(4): 251.

Ren D, Li K, Liu D. 2008. A study of the intellectual ability of 8-14 year-old children in high fluoride, low iodine areas. *Fluoride* 41: 319-320.

Riddell JK, Malin AJ, 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: 105190.

Rocha-Amador D, Navarro ME, Carrizales L, Morales R, Calderon J. 2007. Decreased intelligence in children and exposure to fluoride and arsenic in drinking water. *Cad Saude Publica* 23 Suppl 4: S579-587.

Rocha-Amador D, Navarro M, Trejo-Acevedo A, Carrizales L, Perez-Maldonado I, Diaz-Barriga F, Calderon J. 2009. Use of the Rey-Osterrieth Complex Figure Test for neurotoxicity evaluation of mixtures in children. *Neurotox* 30: 1149-1154.

Rotton J, Tikofsky RS, Feldman HT. 1982. Behavioral effects of chemicals in drinking water. *J Appl Psychol* 67: 230-238.

Russ TC, Killin LOJ, Hannah J, Batty GD, Deary IJ, Starr JM. 2019. Aluminium and fluoride in drinking water in relation to later dementia risk. *Brit J Psychol* 14: 1-6.

Saxena S, Sahay A, Goel P. 2012. Effect of fluoride exposure on the intelligence of school children in Madhya Pradesh, India. *J Neurosci Rural Pract* 3: 144-149.

Sebastian ST, Sunitha S. 2015. A cross-sectional study to assess the intelligence quotient (IQ) of school going children aged 10-12 years in villages of Mysore District, India with different fluoride levels. *J Indian Soc Pedod Prev Dent* 33: 307-311.

Seraj B, Shahrabi M, Falahzade M, Falahzade F, Akhondi N. 2006. Effect of high fluoride concentration in drinking water on children's intelligence. *J Dent Med* 19(2): 80-86.

Seraj B, Shahrabi M, Shadfar M, Ahmadi R, Fallahzadeh M, Eslamlu HF, Kharazifard MJ. 2012. Effect of high water fluoride concentration on the intellectual development of children in Makoo, Iran. *J Dent* 9: 221-229.

Shannon FT, Fergusson DM, Horwood LJ. 1986. Exposure to fluoridated public water supplies and child health and behaviour. *New Zeal Med J* 99: 416-418.

Shao Q. 2003. Study of cognitive function impairment caused by chronic fluorosis. *Chin J Endemiol* 22(4): 336-338.

Sharma JD, Sohu D, Jain P. 2009. Prevalence of neurological manifestations in a human population exposed to fluoride in drinking water. *Fluoride* 42: 127-132.

Shivaprakash PK, Ohri K, Noorani H. 2011. Relation between dental fluorosis and intelligence quotient in school children of Bagalkot District. *J Indian Soc Pedod Prev Dent* 29: 117-120.

Singh A, Jolly SS, Devi P, Bansal BC, Singh SS. 1962. Endemic fluorosis: An epidemiological, biochemical and clinical study in the Bhatinda District of Panjab. *Indian J Med Res* 50: 387-398.

Singh N, Verma KG, Verma P, Sidhu GK, Sachdeva S. 2014. A comparative study of fluoride ingestion levels, serum thyroid hormone & TSH level derangements, dental fluorosis status among school children from endemic and non-endemic fluorosis areas. *Springerplus* 3: 7.

Singh V, Singh C, Sandeep T, Sandeep K, Vikas G, Mukesh T, Anurag T. 2013. A correlation between serum vitamin, acetylcholinesterase activity and IQ in children with excessive endemic fluoride exposure in Rajasthan, India. *Int Res J Medical Sci* 1(3): 12-16.

Soto-Barreras U, Escalante-Villalobos KY, Holguin-Loya B, Perez-Aguirre B, Nevarez-Rascon A, Martinez-Martinez RE, Loyola-Rodriguez JP. 2019. Effect of fluoride in drinking water on dental caries and IQ in children. *Fluoride* 52: 474-482.

Still CN, Kelley P. 1980. The incidence of primary degenerative dementia vs. water fluoride content in South Carolina, USA. *Neurotox* 1: 125-132.

Sudhir KM, Chandu GN, Prashant GM, Subba Reddy VV. 2009. Effect of fluoride exposure on intelligence quotient (IQ) among 13-15 year old school children of known endemic area of fluorosis, Nalgonda District, Andhra Pradesh. *J Indian Assoc Public Health Dent* 2009(13): 88-94.

Sun M, Li S, Wang Y, Li F. 1991. Measurement of intelligence by drawing test among the children in the endemic area of Al-F combined toxicosis. *J Guiyang Med Coll* 16(3): 204-206.

Susheela AK, Bhatnagar M, Vig K, Mondal NK. 2005. Excess fluoride ingestion and thyroid hormone derangements in children living in Delhi, India. *Fluoride* 38: 98-108.

Tamboli BL, Mathur GM, Mathur AP, Lalla SK, Goyal OP. 1980. Prevalence of fluorosis in Pratabpura and Surajpura villages, District Ajmer (Rajasthan). *Indian J Med Res* 71: 57-67.

Till C, Green R, Flora D, Hornung R, Martinez-Mier EA, Blazer M, Farmus L, Ayotte P, Muckle G, Lanphear B. 2020. Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. *Environ Int* 134: 105315.

Tripathi P, Sultana N. 2007. Fluoride content of groundwater and prevalence of dental, skeletal and neurological stage of fluorosis in Tehsil Purwa of Unnao. *Indian J Environ Prot* 27: 737-739.

Trivedi MH, Verma RJ, Chinoy NJ, Patel RS, Sathawara NG. 2007. Effect of high fluoride water on intelligence of school children in India. *Fluoride* 40: 178-183.

Trivedi M, Sangai N, Patel R, Payak M, Vyas S. 2012. Assessment of groundwater quality with special reference to fluoride and its impact on IQ of schoolchildren in six villages of the Mundra Region, Kachchh, Gujurat, India. *Fluoride* 45(4): 377-383.

Valdez Jimenez L, Lopez Guzman OD, Cervantes Flores M, Costilla-Salazar R, Calderon Hernandez J, Alcaraz Contreras Y, Rocha-Amador DO. 2017. In utero exposure to fluoride and cognitive development delay in infants. *Neurotox* 59: 65-70.

Wang A, Duan L, Huang H, Ma J, Zhang Y, Ma Q, Guo Y, Li Z, Cheng X, Zhu J, Zhou G, Ba Y. 2020. Association between fluoride exposure and behavioural outcomes of school-age children: A pilot study in China. *Int J Environ Health Res*: 1-10.

Wang G, Yang D, Jia F, Wang H. 1996. [A study of the IQ levels of four- to seven-year-old children in high fluoride areas]. *Endem Dis Bull* 11(1): 60-62.

Wang G, Yang D, Jia F, Wang H. 2008. A study of the IQ levels of four- to seven-year-old children in high fluoride areas. *Fluoride* 41: 340-343.

Wang G, Gao M, Zhang M, Yang M, Xiang Q. 2012. Correlation between total fluoride intake and children's IQ. *J Southeast Univ Med Ed*: 743-746.

Wang M, Liu L, Li H, Li Y, Liu H, Hou C, Zeng Q, Li P, Zhao Q, Dong L, Zhou G, Yu X, Liu L, Guan Q, Zhang S, Wang A. 2020. Thyroid function, intelligence, and low-moderate fluoride exposure among Chinese school-age children. *Environ Int* 134: 105229.

Wang S, Wang L, Hu P, Guo S, Law S. 2001. [Effects of high iodine and high fluorine on children's intelligence and thyroid function]. *Chin J Endemiol* 20(4): 288-290.

Wang SX, Wang ZH, Cheng XT, Li J, Sang ZP, Zhang XD, Han LL, Qiao XY, Wu ZM, Wang ZQ. 2005. [Investigation and evaluation on intelligence and growth of children in endemic fluorosis and arsenism areas]. *Chin J Endemiol* 24: 179-182.

Wang SX, Wang ZH, Cheng XT, Li J, Sang ZP, Zhang XD, Han LL, Qiao XY, Wu ZM, Wang ZQ. 2007. Arsenic and fluoride exposure in drinking water: Children's IQ and growth in Shanyin County, Shanxi Province, China. *Environ Health Perspect* 115: 643-647.

Wang SY, Zhang HX, Fan W, Fang SJ, Kang PP, Chen XG, Yu MJ. 2005. [The effects of endemic fluoride poisoning caused by coal burning on the physical development and intelligence of children]. *J Appl Clin Ped* 20(9): 897-899.

Wang SY, Zhang HX, Fan W, Fang SJ, Kang PP, Chen XG, Yu MJ. 2008. The effects of endemic fluoride poisoning caused by coal burning on the physical development and intelligence of children. *Fluoride* 41: 344-348.

Wang Z, Wang S, Zhang X, Li J, Zheng X, Hu C. 2006. Investigation of children's growth and development under long-term fluoride exposure. *Chin J Control Endem Dis* 21(4): 239-241.

Wei N, Li Y, Deng J, Xu S, Guan Z. 2014. [The effects of comprehensive control measures on intelligence of school-age children in coal-burning-borne endemic fluorosis areas]. *Chin J Endemiol* 33(2): 320-322.

Xiang Q, Liang Y, Chen L, Wang C, Chen B, Chen X, Zhou M. 2003. Effect of fluoride in drinking water on children's intelligence. *Fluoride* 36: 84-94.

Xiang Q, Liang Y, Chen B, Chen L. 2011. Analysis of children's serum fluoride levels in relation to intelligence scores in a high and low fluoride water village in China. *Fluoride* 44: 191-194.

Xu Y, Lu C, Zhang X. 1994. Effect of fluoride on children's intelligence. *Endem Dis Bull* 2: 83-84.

Yang Y, Wang X, Guo X, Hu P. 1994. [The effects of high levels of fluoride and iodine on child intellectual ability and the metabolism of fluoride and iodine]. *Chin J Epidemiol* 15(4): 296-298.

Yang Y, Wang X, Guo X, Hu P. 2008. The effects of high levels of fluoride and iodine on child intellectual ability and the metabolism of fluoride and iodine. *Fluoride* 41: 336-339.

Yao L, Zhou J, Wang S, Cui K, Lin F. 1996. Analysis of TSH levels and intelligence of children residing in high fluorosis areas. *Lit Inf Prev Med* 2(1): 26-27.

Yao Y. 1997. Comparative assessment of the physical and mental development of children in endemic fluorosis area with water improvement and without water improvement. *Lit Inf Prev Med* 3(1): 42-43.

Yasmin S, Ranjan S, D'Souza D. 2013. Effect of excess fluoride ingestion on human thyroid function in Gaya Region, Bihar, India. *Toxicol Environ Chem* 95: 1235-1243.

Yazdi SM, Sharifian A, Dehghani-Beshne M, Momeni VR, Aminian O. 2011. Effects of fluoride on psychomotor performance and memory of aluminum potroom workers. *Fluoride* 44: 158-162.

Yu YN, Yang WX, Dong Z, Wan CW, Zhang JT, Liu JL, Xiao KQ, Huang YS, Lu BF. 1996. [Neurotransmitter and receptor changes in the brains of fetuses from areas of endemic fluorosis]. *Chin J Endemiol* 15(5): 257-259.

Yu YN, Yang WX, Dong Z, Wan CW, Zhang JT, Liu JL, Xiao KQ, Huang YS, Lu BF. 2008. Neurotransmitter and receptor changes in the brains of fetuses from areas of endemic fluorosis. *Fluoride* 41: 134-138.

Yu X, Chen J, Li Y, Liu H, Hou C, Zeng Q, Cui Y, Zhao L, Li P, Zhou Z, Pang S, Tang S, Tian K, Zhao Q, Dong L, Xu C, Zhang X, Zhang S, Liu L, Wang A. 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 J, Yao H, Chen Y. 1998. [Effect of high level of fluoride and arsenic on children's intelligence]. *Chin J Public Health* 17(2): 57.

Zhang S, Zhang X, Liu H, Qu W, Guan Z, Zeng Q, Jiang C, Gao H, Zhang C, Lei R, Xia T, Wang Z, Yang L, Chen Y, Wu X, Cui Y, Yu L, Wang A. 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: 238-245.

Zhao LB, Liang GH, Zhang DN, Wu XR. 1996. Effect of a high fluoride water supply on children's intelligence. *Fluoride* 29: 190-192.

Zhao Q, Niu Q, Chen J, Xia T, Zhou G, Li P, Dong L, Xu C, Tian Z, Luo C, Liu L, Zhang S, Wang A. 2019. Roles of mitochondrial fission inhibition in developmental fluoride neurotoxicity: Mechanisms of action in vitro and associations with cognition in rats and children. *Arch Toxicol* 93(3): 709-726.

Zhao Q, Tian Z, Zhou G, Niu Q, Chen J, Li P, Dong L, Xia T, Zhang S, Wang A. 2020. SIRT1-dependent mitochondrial biogenesis supports therapeutic effects of resveratrol against neurodevelopment damage by fluoride. *Theranostics* 10: 4822-4838.

Zhou G, Tang S, Yang L, Niu Q, Chen J, Xia T, Wang S, Wang M, Zhao Q, Liu L, Li P, Dong L, Yang K, Zhang S, Wang A. 2019. Effects of long-term fluoride exposure on cognitive ability and the underlying mechanisms: Role of autophagy and its association with apoptosis. *Toxicol Appl Pharmacol 378*: 114608.

Studies Not Available in HAWC

Bachinskii PP, Gutsalenko OA, Naryzhniuk ND, Sidora VD, Shliakhta AI. 1985. [Action of the body fluorine of healthy persons and thyroidopathy patients on the function of hypophyseal-thyroid the system]. *Probl Endokrinol* 31: 25-29.

Balabolkin MI, Mikhailets ND, Lobovskaia RN, Chernousova NV. 1995. [The interrelationship of the thyroid and immune statuses of workers with long-term fluorine exposure]. *Ter Arkh* 67: 41-42.

Baum K, Boerner W, Reiners C, Moll E. 1981. [Bone density and thyroid function in adolescents in relation to fluoride content of drinking water]. *Fortschr Med* 99: 1470-1472.

Berry WTC, Whittles JH. 1963. Absence of effect of fluoride upon the incidence of thyroid enlargements in Wiltshire schoolgirls. *Mon Bull Minist Health Public Health Lab Serv* 22: 50-52.

Cherkinskii SN, Zaslavskaia RM. 1956. [Significance of fluorides in potable water in the development of endemic goiter]. *Probl Endokrinol Gormonoter* 2: 70-75.

Choubisa SL. 2001. Endemic fluorosis in southern Rajasthan, India. *Fluoride* 34: 61-70.

Chuka A, Zhukovskil V, Mirku I, Postel'Niku D. 1964. Prezhdevremennoe starenie naseleniya v zone rasprostraneniya endemicheskogo zoba. *Vestnik Akad Med Nauk Sssr* 19: 23-27.

Dai HX, Zeng P, Wang KY, Zhang XG, Ma ZJ, Zhou YG, Fan ZX, Guo SH. 2013. [Analysis of a survey results of patients with suspected high iodine goiter in Liuji Town Fuping County of Shaanxi Province]. *Chin J Endemiol* 32: 408-411.

Day T, Powell-Jackson P. 1972. Fluoride, water hardness, and endemic goitre. *Lancet* 299(7761): 1135-1138.

Desai VK, Solanki DM, Bansal RK. 1993. Epidemiological study of goitre in endemic fluorosis district of Gujarat. *Fluoride* 26: 187-190.

Díaz-Cadórniga FJ, Delgado E, Tartón T, Valdés MM, Méndez A, Fernández MT, Rojo C. 2003. Endemic goiter associated with high iodine intake in primary school children in the Saharawi Arab Democratic Republic. *Endocrinol Nutr* 50: 357-362.

Eichner R, Borner W, Henschler D, Kohler W, Moll E. 1981. [Osteoporosis therapy and thyroid function. Influence of 6 months of sodium fluoride treatment on thyroid function and bone density]. *Fortschr Med* 99: 342-348.

Fiorentini S, Galeazzi M, Visintin B. 1947. II fluoro in natura come agente morbigeno II. La fluorosi die Campagnano di Roma. III. Un focolaio di fluorosi umana a Campagnano di Roma. IV. Osservazioni radiologiche sui processi alveolari, sulle ossa mascellari, e sul paradenzio degli abitanti die Campagnano. V. Zona fluorotica intorno a Campagnano di Roma. VI. Frequenza e caratteri clinici della carie dentale in soggetti fluorotici. *Rend Ist Superiore Sanita* 10: 721-804.

Gabovich RD, Verzhikovskaia NV. 1958. [Effect of fluoride compounds on absorption of radioactive iodine by the thyroid gland in humans and in experimental conditions]. *Probl Endokrinol Gormonoter* 4: 49-54.

Galletti P, Held HR, Korrodi H, Wegmann T. 1956. [Investigations on the possible damaging effect of fluorine on the thyroid gland]. *Helv Med Acta* 23: 601-605.

Galletti PM, Joyet G, Jallut O. 1957. [Effect of sodium fluoride on thyroid function in Basedow's Disease]. *Helv Med Acta* 24: 209-215.

Galletti PM, Joyet G. 1958. Effect of fluorine on thyroidal iodine metabolism in hyperthyroidism. *J Clin Endocrinol Metab* 18: 1102-1110.

Gas'kov AI, Savchenkov MF, Iushkov NN. 2005. [The specific features of the development of iodine deficiencies in children living under environmental pollution with fluorine compounds]. *Gig Sanit*: 53-55.

Gedalia I, Brand N. 1963. The relationship of fluoride and iodine in drinking water in the occurrence of goiter. *Arch Int Pharmacodyn Ther* 142: 312-315.

Grimm H. 1973. [The physical development of schoolchildren under the influence of drinking water fluoridation in Karl Marx Stadt]. *Dtsch Gesundheitsw* 28: 2363-2369.

Hasling C, Nielsen HE, Melsen F, Mosekilde L. 1987. Safety of osteoporosis treatment with sodium fluoride, calcium phosphate and vitamin D. *Miner Electrolyte Metab* 13: 96-103.

Hidehiko T. 1958. On the relation between the distribution of endemic goiter and the fluorine content of natural water in Hidaka Province, Hokkaido. *Folia Pharmacol Jpn* 54: 225-229.

Hoffmann-Axthelm W. 1953. [Observations on the influence of fluorine on dental enamel and thyroid gland]. *Dtsch Zahnarztl Z* 8: 757-765.

Jentzer A. 1956. [Effect of fluorine on the iodine content of the human thyroid gland]. *Bull Schweiz Akad Med Wiss* 12: 539-543.

Jooste PL, Weight MJ, Kriek JA, Louw AJ. 1999. Endemic goitre in the absence of iodine deficiency in schoolchildren of the Northern Cape Province of South Africa. *Eur J Clin Nutr* 53: 8-12.

Kolomiitseva MG. 1961. [The content of fluorine in the external environment of the Upper Altai autonomous region and its role in the etiology of endemic goiter]. *Gig Sanit* 26: 101-103.

Korrodi H, Wegmann T, Galleti P, Held HR. 1955. [Caries prophylaxis and the untoward effects of fluor on the thyroid gland]. *Schweiz Med Wochenschr* 85: 1016-1019.

Kutlucan A, Kale Koroglu B, Numan Tamer M, Aydin Y, Baltaci D, Akdogan M, Ozturk M, Vural H, Ermis F. 2013. The investigation of effects of fluorosis on thyroid volume in school-age children. *Med Glas* 10: 93-98.

Latham MC, Grech P. 1967. The effects of excessive fluoride intake. *Am J Public Health* 57: 651-660.

Leone NC, Leatherwood EC, Petrie IM, Lieberman L. 1964. Effect of fluoride on thyroid gland: Clinical study. *J Am Dent Assoc* 69: 179-180.

Levi JE, Silberstein HE. 1955. Lack of effect of fluorine ingestion on uptake of iodine 131 by the thyroid gland. *J Lab Clin Med* 45: 348-351.

McGlashan N, Chelkowska E, Sasananan S. 2010. A survey of goiter morbidity in Ban Mae Toen, northwest Thailand. *Southeast Asian J Trop Med Public Health* 41: 1200-1208.

Rathore S, Meena C, Gonmei Z, Dwivedi S, Toteja GS, Bala K. 2018. Study of excess fluoride ingestion and thyroid hormone derangement in relation with different fluoride levels in drinking water among children of Jodhpur District, Rajasthan, India. *Asian J Microbiol Biotechnol Environ Sci* 20: 327-331.

Reisenauer R, Rezler D, Křemenová J, Preininger Q. 1961. [Fluorization of the waters in Czechoslovakia. IV. Endocrinological control of results of two years' fluorization of drinking-water in school children]. *Cesk Stomatol* 61: 91-97.

Romer TEZ, Kowalczyk B, Kacprzak M, Wiktorowski M. 1976. [Incidence of goiter in pubertal girls of the Piotrkow Region and iodide content in drinking water]. *Endokrynol Pol* 27: 373-380.

Savchenkov MF, Efimova NV, Manueva RS, Nikolaeva LA, Shin NS. 2016. [Thyroid gland pathology in children population exposed to the combination of iodine deficiency and fluoride pollution of environment]. *Gig Sanit* 95: 1201-1205.

Shtifanova AK. 1962. [The fluorine content in water, soil and vegetal products of the Alma-Atinsk District areas and its role in the etiology of dental caries and endemic goiter]. *Zdravookhranenie Kazakhstana*: 60-63.

Siddiqui AH. 1969. Incidence of simple goiter in areas of endemic fluorosis in Nalgonda District, Andhra Pradesh, India. *Fluoride* 2: 200-205.

Sidora VD, Shliakhta AI, Iugov VK, Kas'ianenko AS, Piatenko VG. 1983. [Indices of the pituitary-thyroid system in residents of cities with various fluorine concentrations in drinking water]. *Probl Endokrinol* 29: 32-35.

Sung FC, Chen KP, Chen CY, Tai PW, Yang CF. 1973. Studies of the effect of salt iodization on endemic goiter in Taiwan. IV. A survey of drinking water in relation to endemic goiter. *J Fomosan Med Assoc* 72: 96-103.

Tokar VI, Voroshnin VV, Sherbakov SV. 1989. [Chronic effects of fluorides on the pituitary-thyroid system in industrial workers]. *Gig Tr Prof Zabol*: 19-22.

Wespi HJ. 1954. [Iodized-fluoridized salt for the prevention of goiter and caries]. *Schweiz Med Wochenschr* 84: 885-890.

Yu YN. 1985. [Effects of chronic fluorosis on the thyroid gland]. *Chin Med J* 65: 747-7479.

Studies in Non-human Animals

As described in **Figure 4**, 339 non-human mammal studies were included; however, full data extraction was only conducted on studies with primary neurological outcomes and/or secondary functional neurological outcomes (e.g., motor activity). Data extraction was completed using HAWC. Data were extracted from a subset of included studies in animals (n = 123) and are available in HAWC based on outcome. The following lists of references are organized as studies that are available in HAWC followed by studies that are not available in HAWC. Specifically, all primary outcomes and functional neurological secondary outcomes (e.g., motor activity) were extracted from animal studies and are available in HAWC, including studies from the NTP (2016) assessment. Studies are also available in HAWC that evaluated mechanistic effects related to oral fluoride exposure at or below 20 ppm fluoride drinking water equivalents for categories of mechanistic endpoints with the largest amount of available data (i.e., biochemistry of the brain or neurons, neurotransmission, oxidative stress, and histopathology [n = 70]); however, these mechanistic data were generally not extracted. Several animal studies assessed primary neurological outcomes and/or functional neurological secondary outcomes and mechanistic effects in the four mechanistic categories listed above (n = 56). In total, 140 animal studies are available in HAWC (70 with primary neurological outcomes and/or secondary functional neurological outcomes without relevant mechanistic data; 15 with relevant mechanistic data only; and 55 with primary/or secondary functional neurological outcomes with relevant mechanistic data). Studies that evaluated other mechanistic endpoints, as well as studies that only assessed mechanistic effects at fluoride levels above 20 ppm fluoride drinking water equivalents, are not available in HAWC (n = 199). The list below presents the 339 non-human animal studies that were included in the review. An overview of the screening results is outlined in the study selection diagram (**Figure 4**) that reports numbers of included studies as well as numbers of studies excluded for each reason at the full text review stage.

Studies Available in HAWC

Adedara IA, Abolaji AO, Idris UF, Olabiyi BF, Onibiyo EM, Ojuade TD, Farombi EO. 2017. Neuroprotective influence of taurine on fluoride-induced biochemical and behavioral deficits in rats. *Chem Biol Interact* 261: 1-10.

Adedara IA, Olabiyi BF, Ojuade TD, Idris UF, Onibiyo EM, Farombi EO. 2017. Taurine reverses sodium fluoride-mediated increase in inflammation, caspase-3 activity, and oxidative damage along the brain-pituitary-gonadal axis in male rats. *Can J Phys Pharmacol* 95: 1019-1029.

Agustina F, Sofro ZM, Partadiredja G. 2018. Subchronic administration of high-dose sodium fluoride causes deficits in cerebellar purkinje cells but not motor coordination of rats. *Biol Trace Elem Res* 188(2): 424-433.

Ahmad KR, Noor S, Jabeen S, Nauroze T, Kanwal MA, Raees K, Abbas T. 2017. Amelioration by jambul fruit extract of fluoride-induced hepato-nephronal histopathologies and impaired neuromotor capacity in mice. *Fluoride* 50: 2-14.

Akinrinade ID, Memudu AE, Ogundele OM. 2015. Fluoride and aluminium disturb neuronal morphology, transport functions, cholinesterase, lysosomal and cell cycle activities. *Pathophysiology* 22: 105-115.

Akinrinade ID, Memudu AE, Ogundele OM, Ajetunmobi OI. 2015. Interplay of glia activation and oxidative stress formation in fluoride and aluminium exposure. *Pathophysiology* 22: 39-48.

Baba NA, Raina R, Verma PK, Sultana M. 2014. Alterations in plasma and tissue acetylcholinesterase activity following repeated oral exposure of chlorpyrifos alone and in conjunction with fluoride in Wistar rats. *Proc Indian Natl Sci Acad B Biol Sci* 84: 969-972.

Bagmut I, Kolisnyk I, Titkova A, Babiy L, Filipchenko S. 2018. The antioxidant system enzymes' activity in rats' brain, intoxicated with sodium fluoride in subtoxic doses. *Arch Balkan Med Union* 53(4): 506-511.

Balaji B, Kumar EP, Kumar A. 2015. Evaluation of standardized bacopa monniera extract in sodium fluoride-induced behavioural, biochemical, and histopathological alterations in mice. *Toxicol Ind Health* 31: 18-30.

Balayssac D, Richard D, Authier N, Nicolay A, Jourdan D, Eschalier A, Coudore F. 2002. Absence of painful neuropathy after chronic oral fluoride intake in Sprague-Dawley and Lou/C rats. *Neurosci Lett* 327: 169-172.

Banala RR, Karnati PR. 2015. Vitamin A deficiency: An oxidative stress marker in sodium fluoride (NaF) induced oxidative damage in developing rat brain. *Int J Dev Neurosci* 47: 298-303.

Banala R, Nagapuri K, Mohd K, Reddy M, Karnati P. 2018. Carica Papaya leaf extract as a neuroprotective agent against behavioral and neurotransmitter changes in brain of the rat treated with sodium fluoride in pre- and post-natal periods. *Pharmacogn Mag* 14(55): 123-131.

Banji D, Banji OJ, Pratusha NG, Annamalai AR. 2013. Investigation on the role of spirulina platensis in ameliorating behavioural changes, thyroid dysfunction and oxidative stress in offspring of pregnant rats exposed to fluoride. *Food Chem* 140: 321-331.

Baran-Poesina V, Negres S, Dobrescu D, Dimcevici-Poesina N, Dimcevici-Poesina A, Feghiu A, Soare T, Militaru M. 2013. Experimental pharmacological researches regarding the influence of sodium fluoride in allopathic and homeopathic doses on central nervous system's performances: A correlation between behavioral response in classic maze test and morphological aspects of cerebral cortex. *Farmacia* 61: 781-799.

Bartos M, Gumilar F, Bras C, Gallegos CE, Giannuzzi L, Cancela LM, Minetti A. 2015. Neurobehavioural effects of exposure to fluoride in the earliest stages of rat development. *Physiol Behav* 147: 205-212.

Bartos M, Gumilar F, Gallegos CE, Bras C, Dominguez S, Monaco N, Esandi MDC, Bouzat C, Cancela LM, Minetti A. 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.

Basha PM, Rai P, Begum S. 2011. Fluoride toxicity and status of serum thyroid hormones, brain histopathology, and learning memory in rats: A multigenerational assessment. *Biol Trace Elem Res* 144: 1083-1094.

Basha PM, Sujitha NS. 2012. Combined impact of exercise and temperature in learning and memory performance of fluoride toxicated rats. *Biol Trace Elem Res* 150: 306-313.

Bataineh HN, Nusier MK. 2006. Impact of 12-week ingestion of sodium fluoride on aggression, sexual behavior, and fertility in adult male rats. *Fluoride* 39: 293-301.

Bera I, Sabatini R, Auteri P, Flace P, Sisto G, Montagnani M, Potenza MA, Marasciulo FL, Carratu MR, Coluccia A, Borracci P, Tarullo A, Cagiano R. 2007. Neurofunctional effects of developmental sodium fluoride exposure in rats. *Eur Rev Med Pharmacol Sci* 11: 211-224.

Bhatnagar M, Rao P, Sushma J, Bhatnagar R. 2002. Neurotoxicity of fluoride: Neurodegeneration in hippocampus of female mice. *Indian J Exp Biol* 40: 546-554.

Bhatnagar M, Sukhwal P, Suhalka P, Jain A, Joshi C, Sharma D. 2011. Effects of fluoride in drinking water on NADPH-diaphorase neurons in the forebrain of mice: A possible mechanism of fluoride neurotoxicity. *Fluoride* 44: 195-209.

Chen H, Geng D. 2011. [The change of cognition induced by chronic fluoride in rats]. *Acta Academiae Medicinae Xuzhou* 31(5): 319-322.

Chen J, Niu Q, Xia T, Zhou G, Li P, Zhao Q, Xu C, Dong L, Zhang S, Wang A. 2018. ERK1/2-mediated disruption of BDNF-TrkB signaling causes synaptic impairment contributing to fluoride-induced developmental neurotoxicity. *Toxicology* 410: 222-230.

Chinoy NJ, Shah SD. 2004. Biochemical effects of sodium fluoride and arsenic trioxide toxicity and their reversal in the brain of mice. *Fluoride* 37: 80-87.

Chioca LR, Raupp IM, Da Cunha C, Losso EM, Andreatini R. 2008. Subchronic fluoride intake induces impairment in habituation and active avoidance tasks in rats. *Eur J Pharmacol* 579: 196-201.

Chouhan S, Flora SJS. 2008. Effects of fluoride on the tissue oxidative stress and apoptosis in rats: Biochemical assays supported by IR spectroscopy data. *Toxicology* 254: 61-67.

Chouhan S, Lomash V, Flora SJ. 2010. Fluoride-induced changes in haem biosynthesis pathway, neurological variables and tissue histopathology of rats. *J Appl Toxicol* 30: 63-73.

Cui YS, Zhong Q, Li WF, Liu ZH, Wang Y, Hou CC. 2017. [Effects of fluoride exposure on thyroid hormone level and intelligence in rats]. *Chin J Ind Hyg Occup Dis* 35: 888-892.

Dabrowska E. 1997. Effect of different fluorine doses on the supraoptic nucleus of the rat. *Folia Histochem Cytobiol* 35: 115-116.

Dong Y, Wang Y, Wei N, Guan Z. 2015. [Expression levels of brain muscarinic acetylcholine receptor in offspring rats of drinking-water borne fluorosis]. *Chin J Endemiol* 34: 326-330.

Dong YT, Wang Y, Wei N, Zhang QF, Guan ZZ. 2015. Deficit in learning and memory of rats with chronic fluorosis correlates with the decreased expressions of M1 and M3 muscarinic acetylcholine receptors. *Arch Toxicol* 89: 1981-1991.

Dong YT, Wei N, Qi XL, Liu XH, Chen D, Zeng XX, Guan ZZ. 2017. Attenuating effect of vitamin E on the deficit of learning and memory of rats with chronic fluorosis: The mechanism may involve muscarinic acetylcholine receptors. *Fluoride* 50: 354-364.

Dong YW, Y. Wei, N. Guan, Z. 2015. [Expression of muscarinic acetylcholine receptors in the brain of rats with chronic fluorosis]. *Chin J Endemiol* 34(2): 84-88.

Ekambaram P, Paul V. 2001. Calcium preventing locomotor behavioral and dental toxicities of fluoride by decreasing serum fluoride level in rats. *Environ Toxicol Pharmacol* 9: 141-146.

Ekambaram P, Paul V. 2002. Modulation of fluoride toxicity in rats by calcium carbonate and by withdrawal of fluoride exposure. *Pharmacol Toxicol* 90: 53-58.

Ekambaram P, Paul V. 2003. Effect of vitamin D on chronic behavioral and dental toxicities of sodium fluoride in rats. *Fluoride* 36: 189-197.

El-lethey HS, Kamel MM, Shaheed IB. 2010. Neurobehavioral toxicity produced by sodium fluoride in drinking water of laboratory rats. *J Am Sci* 6(5): 54-63.

El-lethey HS, Kamel MM. 2011. Effects of black tea in mitigation of sodium fluoride potency to suppress motor activity and coordination in laboratory rats. *J Am Sci* 7(4): 243-254.

El-lethey HS, Shaheed IB. 2011. Potential health impact of black tea against Na-F-induced alterations in territorial aggression, sexual behaviour and fertility of male rats. *Life Sci J* 8: 828-839.

Elliott L. 1967. Lack of effect of administration of fluoride on the central nervous system of rats. *Acta Pharmacol Toxicol (Copenh)* 25: 323-328.

Flace P, Benagiano V, Vermesan D, Sabatini R, Inchingolo AM, Auteri P, Ambrosi G, Tarullo A, Cagiano R. 2010. Effects of developmental fluoride exposure on rat ultrasonic vocalization, acoustic startle reflex and pre-pulse inhibition. *Eur Rev Med Pharmacol Sci* 14: 507-512.

Gabovich RD. 1962. [On the problem of the effect of fluorine in drinking water on the functional state of the central nervous system]. *Gig Sanit* 27: 10-12.

Gao Q, Liu YJ, Wu CX, Long YG, Guan ZZ. 2008. [Level of oxidative stress in rat brains and learning and memory function of rats with chronic fluorosis]. *Chin J Endemiol* 27: 371-373.

Gao Q, Liu YJ, Wu CX, Long YG, Guan ZZ. 2008. [Effects of fluoride on learning and memory and cholinesterase activity in rat brains]. *Chin J Endemiol* 27: 128-130.

Gao Q, Liu YJ, Guan ZZ. 2009. Decreased learning and memory ability in rats with fluorosis: Increased oxidative stress and reduced cholinesterase activity in the brain. *Fluoride* 42: 277-285.

Gao Y, Liu L, Young L, Huan L, Jin H. 2009. Effects of learning and memory of fluoride and the antagonism of selenium in rats. *Studies of Trace Elements and Health* 26(2): 1-3.

Ge QD, Tan Y, Luo Y, Wang WJ, Zhang H, Xie C. 2018. MiR-132, miR-204 and BDNF-TrkB signaling pathway may be involved in spatial learning and memory impairment of the offspring rats caused by fluorine and aluminum exposure during the embryonic stage and into adulthood. *Environ Toxicol Pharmacol* 63: 60-68.

Ge Y, Chen L, Yin Z, Song X, Ruan T, Hua L, Liu J, Wang J, Ning H. 2018. Fluoride-induced alterations of synapse-related proteins in the cerebral cortex of ICR offspring mouse brain. *Chemosphere* 201: 874-883.

Gopal K, Saxena R, Gupta GSD, Rana MD, Agrawal D. 2006. Fluoride induced alterations in neurobehavioural and cardiovascular responses in rats. *J Adv Zool* 27: 1-7.

Gui CZ, Ran LY, Wu CX, Long YG, He J, Zhang H, Guan ZZ. 2009. [Changes in learning and memory ability and brain cholinesterase activity in the rats with coal burning fluorosis]. *Chin J Endemiol* 28: 497-500.

Gui CZ, Ran LY, Li JP, Guan ZZ. 2010. Changes of learning and memory ability and brain nicotinic receptors of rat offspring with coal burning fluorosis. *Neurotoxicol Teratol* 32: 536-541.

Gui CZ, Ran LY, Guan ZZ. 2011. [Expression levels of brain nicotinic acetylcholine receptor mRNA and protein in coal-burning type of fluorosis rats]. *Chin J Endemiol* 30: 239-242.

Guner S, Uyar-Bozkurt S, Haznedaroglu E, Mentes A. 2016. Dental fluorosis and catalase immunoreactivity of the brain tissues in rats exposed to high fluoride pre- and postnatally. *Biol Trace Elem Res* 174: 150-157.

Han H, Du W, Zhou B, Zhang W, Xu G, Niu R, Sun Z. 2014. Effects of chronic fluoride exposure on object recognition memory and mRNA expression of SNARE complex in hippocampus of male mice. *Biol Trace Elem Res* 158: 58-64.

Hong JH, Ge YM, Ning HM, Wang JD. 2005. [Effects of High Fluoride and Low Iodine on Learning-Memory and TchE of Brain in Offspring Rats]. *Chin Prev Med* 6: 489-491.

Inkielewicz I, Krechniak J. 2004. Fluoride effects on glutathione peroxidase and lipid peroxidation in rats. *Fluoride* 37: 7-12.

Jain A, Mehta VK, Chittora RA, Mahdi A, Bhatnagar M. 2015. Melatonin ameliorates fluoride induced neurotoxicity in young rats: An in vivo evidence. *Asian J Pharm Clin Res* 8: 164-167.

Jetti R, Raghuveer CV, Mallikarjuna RC. 2016. Protective effect of ascorbic acid and Ginkgo biloba against learning and memory deficits caused by fluoride. *Toxicol Ind Health* 32: 183-187.

Jia B, Zong L, Lee JY, Lei J, Zhu Y, Xie H, Clemens JL, Feller MC, Na Q, Dong J, McLane MW, Jones-Beatty K, Burd I. 2019. Maternal supplementation of low dose fluoride alleviates adverse perinatal outcomes following exposure to intrauterine inflammation. *Sci Rep* 9(1): 2575.

Jiang C, Zhang S, Liu H, Guan Z, Zeng Q, Zhang C, Lei R, Xia T, Wang Z, Yang L, Chen Y, Wu X, Zhang X, Cui Y, Yu L, Wang A. 2014. Low glucose utilization and neurodegenerative changes caused by sodium fluoride exposure in rat's developmental brain. *Neuromolecular Med* 16: 94-105.

Jiang S, Su J, Yao S, Zhang Y, Cao F, Wang F, Wang H, Li J, Xi S. 2014. Fluoride and arsenic exposure impairs learning and memory and decreases mGluR5 expression in the hippocampus and cortex in rats. *PLoS One* 9(4): e96041.

Khan AM, Raina R, Dubey N, Verma PK. 2017. Effect of deltamethrin and fluoride co-exposure on the brain antioxidant status and cholinesterase activity in Wistar rats. *Drug Chem Toxicol* 41(2): 1-5.

Kinawy AA, Al-Eidan AA. 2018. Impact of prenatal and postnatal treatment of sodium fluoride and aluminum chloride on some hormonal and sensorimotor aspects in rats. *Biol Trace Elem Res*: 1-8.

Kivrak Y. 2012. Effects of fluoride on anxiety and depression in mice. *Fluoride* 45: 302-306.

Li M, Cui J, Gao YH, Zhang W, Sun LY, Liu XN, Liu Y, Sun DJ. 2015. Pathological changes and effect on the learning and memory ability in rats exposed to fluoride and aluminum. *Toxicol Res* 4: 1366-1373.

Li X, Zhang J, Niu R, Manthari RK, Yang K, Wang J. 2019. Effect of fluoride exposure on anxiety- and depression-like behavior in mouse. *Chemosphere* 215: 454-460.

Liu F, Ma J, Zhang H, Liu P, Liu YP, Xing B, Dang YH. 2014. Fluoride exposure during development affects both cognition and emotion in mice. *Physiol Behav* 124: 1-7.

Liu WX. 1989. [Experimental study of behavior and cerebral morphology of rat pups generated by fluorotic female rat]. *Chin J Pathol* 18: 290-292.

Liu YJ, Gao Q, Wu CX, Long YG, Guan ZZ. 2009. [Modified expression of extracellular signal-regulated protein kinase signal transduction in rat brains and changed capacity of learning and memory of rats with chronic fluorosis]. *Chin J Endemiol* 28: 32-35.

Liu YJ, Gao Q, Wu CX, Guan ZZ. 2010. Alterations of nAChRs and ERK1/2 in the brains of rats with chronic fluorosis and their connections with the decreased capacity of learning and memory. *Toxicol Lett* 192: 324-329.

Liu YJ, Gao Q, Long YG, Yu YN, Guan ZZ. 2011. [Influence of chronic fluorosis on expression of phospho-Elk-1 in rat brains]. *Chin J Endemiol* 30: 251-255.

Lou DD, Guan ZZ, Liu YJ, Liu YF, Zhang KL, Pan JG, Pei JJ. 2013. The influence of chronic fluorosis on mitochondrial dynamics morphology and distribution in cortical neurons of the rat brain. *Arch Toxicol* 87: 449-457.

Lu F, Zhang Y, Trivedi A, Jiang X, Chandra D, Zheng J, Nakano Y, Uyghurturk DA, Jalai R, Guner S, Mentes A, DenBesten PK. 2019. Fluoride related changes in behavioral outcomes may relate to increased serotonin. *Physiol Behav* 206: 76-83.

Ma J, Liu F, Liu P, Dong YY, Chu Z, Hou TZ, Dang YH. 2015. Impact of early developmental fluoride exposure on the peripheral pain sensitivity in mice. *Int J Dev Neurosci* 47: 165-171.

Manusha S, Sudhakar K, Reddy KP. 2019. Protective effects of allium sativum extract against sodium fluoride induced neurotoxicity. *Int J Pharm Sci Res* 10(2): 625-633.

McPherson CA, Zhang G, Gilliam R, Brar SS, Wilson R, Brix A, Picut C, Harry GJ. 2018. An evaluation of neurotoxicity following fluoride exposure from gestational through adult ages in Long-Evans hooded rats. *Neurotoxicol Res*: 1-18.

Mesram N, Nagapuri K, Banala RR, Nalagoni CR, Karnati PR. 2016. Quercetin treatment against NaF induced oxidative stress related neuronal and learning changes in developing rats. *J King Saud Univ Sci* 29: 221-229.

Mullenix PJ, Denbesten PK, Schunior A, Kernan WJ. 1995. Neurotoxicity of sodium fluoride in rats. *Neurotoxicol Teratol* 17: 169-177.

Nageshwar M, Sudhakar K, Reddy NCC, Reddy KP. 2017. Neuroprotective effects of curcumin on sodium fluoride induced behavioural and enzymatic changes in brain and muscles of rat. *J Environ Biol* 38: 675-681.

Nageshwar M, Sudhakar K, Reddy KP. 2018. Quercetin ameliorates oxidative stress, neural damage of brain and behavioral impairment of rat with fluoride exposure. *Int J Pharm Sci Res* 9(8): 3247-3256.

Nian W, Wang X, Shao D, Yu Q, Ouyang W, Zhang Z, Ruan Q. 2018. Effects of subchronic exposure to fluorine on hippocampal injury in mice and its molecular mechanism. *Acta Sci Circumst* 38(11): 4512-4519.

Niu Q, Chen J, Xia T, Li P, Zhou G, Xu C, Zhao Q, Dong L, Zhang S, Wang A. 2018. Excessive ER stress and the resulting autophagic flux dysfunction contribute to fluoride-induced neurotoxicity. *Environ Pollut* 233: 889-899.

Niu R, Sun Z, Wang J, Cheng Z. 2008. Effects of fluoride and lead on locomotor behavior and expression of nissl body in brain of adult rats. *Fluoride* 41: 276-282.

Niu R, Sun Z, Cheng Z, Li Z, Wang J. 2009. Decreased learning ability and low hippocampus glutamate in offspring rats exposed to fluoride and lead. *Environ Toxicol Pharmacol* 28: 254-258.

Niu R, Liu S, Wang J, Zhang J, Sun Z. 2014. Proteomic analysis of hippocampus in offspring male mice exposed to fluoride and lead. *Biol Trace Elem Res* 162: 227-233.

Niu R, Xue X, Zhao Y, Sun Z, Yan X, Li X, Feng C, Wang J. 2015. Effects of fluoride on microtubule ultrastructure and expression of Tubalpha1a and Tubbeta2a in mouse hippocampus. *Chemosphere* 139: 422-427.

Niu R, Chen H, Manthari RK, Sun Z, Wang J, Zhang J, Wang J. 2018. Effects of fluoride on synapse morphology and myelin damage in mouse hippocampus. *Chemosphere* 194: 628-633.

Nkpaa KW, Onyeso GI. 2018. Rutin attenuates neurobehavioral deficits, oxidative stress, neuro-inflammation and apoptosis in fluoride treated rats. *Neurosci Lett* 682: 92-99.

Paul V, Ekambaram P, Jayakumar AR. 1998. Effects of sodium fluoride on locomotor behavior and a few biochemical parameters in rats. *Environ Toxicol Pharmacol* 6: 187-191.

Pereira M, Dombrowski PA, Losso EM, Chioca LR, Da Cunha C, Andreatini R. 2011. Memory impairment induced by sodium fluoride is associated with changes in brain monoamine levels. *Neurotoxicol Res* 19: 55-62.

Pulungan ZS, Sofro ZM, Partadiredja G. 2016. Sodium fluoride does not affect the working memory and number of pyramidal cells in rat medial prefrontal cortex. *Anat Sci Int* 93(1): 128-138.

Raghu J, Raghuveer VC, Rao MC, Somayaji NS, Babu PB. 2013. The ameliorative effect of ascorbic acid and Ginkgo biloba on learning and memory deficits associated with fluoride exposure. *Interdiscip Toxicol* 6: 217-221.

Raju S, Sivanesan SK, Gudemalla K. 2019. Cognitive enhancement effect of Ginkgo biloba extract on memory and learning impairments induced by fluoride neurotoxicity. *Int J Res Pharm Sci* 10(1): 129-134.

Reddy MM, Karnati PR. 2015. Protective effects of aqueous extract of fruit pulp of tamarindus indica on motor activity and metabolism of the gastrocnemius muscle of rats treated with fluoride. *Int J Toxicol Pharmacol Res* 7: 241-246.

Reddy YP, Tiwari SK, Shaik AP, Alsaeed A, Sultana A, Reddy PK. 2014. Effect of sodium fluoride on neuroimmunological parameters, oxidative stress and antioxidative defenses. *Toxicol Mech Methods* 24: 31-36.

Rumiantsev GI, Novikov SM, Mel'nikova NN, Levchenko NI, Kozeeva EE. 1988. [Experimental study of the biological effect of salts of hydrofluosilicic acid]. *Gig Sanit*: 80-82.

Sarkozi K, Horvath E, Vezer T, Papp A, Paulik E. 2015. Behavioral and general effects of subacute oral arsenic exposure in rats with and without fluoride. *Int J Environ Health Res* 25: 418-431.

Shah SD, Chinoy NJ. 2004. Adverse effects of fluoride and/or arsenic on the cerebral hemisphere of mice and recovery by some antidotes. *Fluoride* 37: 162-171.

Shalini B, Sharma JD. 2015. Beneficial effects of emblica officinalis on fluoride-induced toxicity on brain biochemical indexes and learning-memory in rats. *Toxicol Int* 22: 35-39.

Shan KR, Qi XL, Long YG, Nordberg A, Guan ZZ. 2004. Decreased nicotinic receptors in PC12 cells and rat brains influenced by fluoride toxicity: A mechanism relating to a damage at the level in post-transcription of the receptor genes. *Toxicology* 200: 169-177.

Sharma C, Suhalka P, Bhatnagar M. 2018. Curcumin and resveratrol rescue cortical-hippocampal system from chronic fluoride-induced neurodegeneration and enhance memory retrieval. *Int J Neurosci*: 1-15.

Shen X, Zhang Z, Xu X. 2004. [Effect of iodine and selenium on learning memory impairment induced by fluorosis and blood biochemical criterion of rats]. *Occupation and Health* 20(1): 6-8.

Sudhakar K, Nageshwar M, Pratap Reddy K. 2017. Seed extract of Abelmoschus moschatus medik reverses NAF-induced behavioral changes through neurodegeneration and oxidative stress in brain of rat. *Asian J Pharm Clin Res* 10: 165-171.

Sudhakar K, Nageshwar M, Reddy KP. 2018. Protective effect of okra, Abelmoschus moschatus seed extract on developing brain of rats during pre- and post-natal fluoride exposure. *Int J Pharm Sci Res* 9: 1519-1528.

Sudhakar K, Nageshwar M, Reddy KP. 2018. Abelmoschus moschatus extract reverses altered pain and neurohistology of a rat with developmental exposure of fluoride. *J Appl Pharm Sci* 8(6): 94-104.

Sudhakar K, Reddy KP. 2018. Protective effects of Abelmoschus moschatus seed extract on neurotransmitter system of developing brain of Wistar rats with gestational and post-natal exposure of sodium fluoride. *Int J Green Pharm* 12: S131-S139.

Sun ZR, Liu FZ, Wu LN, Lu Y, Yu DK. 2000. [Effects of high fluoride drinking water on the cerebral function of mice]. *Chin J Epidemiol* 19: 262-263.

Sun ZR, Liu FZ, Wu LN, Lu Y, Yu DK. 2008. Effects of high fluoride drinking water on the cerebral function of mice. *Fluoride* 41: 148-151.

Sun Z, Zhang Y, Xue X, Niu R, Wang J. 2018. Maternal fluoride exposure during gestation and lactation decreased learning and memory ability, and glutamate receptor mRNA expressions of mouse pups. *Hum Exp Toxicol* 37: 87-93.

Trivedi MH, Verma RJ, Chinoy NJ. 2007. Amelioration by black tea of sodium fluoride-induced changes in protein content of cerebral hemisphere, cerebellum and medulla oblongata in brain region of mice. *Acta Poloniae Pharm* 64: 221-225.

Trivedi MH, Verma RJ, Chinoy NJ. 2009. Mitigation of sodium fluoride induced toxicity in mice brain by black tea infusion. *Fluoride* 42: 29-33.

Trivedi MH, Verma RJ, Sangai NP, Chinoy NJ. 2011. Black tea extract mitigation of NaF-induced lipid peroxidation in different regions of mice brains. *Fluoride* 44: 243-254.

Trivedi MH, Verma RJ, Sangai NP, Chinoy NJ. 2012. Mitigation by black tea extract of sodium fluoride induced histopathological changes in brain of mice. *Fluoride* 45: 13-26.

Tsunoda M, Aizawa Y, Nakano K, Liu Y, Horiuchi T, Itai K, Tsunoda H. 2005. Changes in fluoride levels in the liver, kidney, and brain and in neurotransmitters of mice after subacute administration of fluoride. *Fluoride* 38: 284-292.

Varner JA, Jensen KF, Horvath W, Isaacson RL. 1998. Chronic administration of aluminum-fluoride or sodium-fluoride to rats in drinking water: Alterations in neuronal and cerebrovascular integrity. *Brain Res* 784(1-2): 284-298.

Verma RJ, Trivedi MH, Chinoy NJ. 2007. Black tea amelioration of sodium fluoride-induced alterations of DNA, RNA, and protein contents in the cerebral hemisphere, cerebellum, and medulla oblongata regions of mouse brain. *Fluoride* 40: 7-12.

Wang G, Li J, Zhu H, Zhu J. 2006. Effect of different doses of chronic exposure of fluoride on rat learning and memory behavior. *Studies of Trace Elements and Health* 23(2): 1-2.

Wang JD, Ge YM, Ning HM, Wang SL. 2004. Effects of high fluoride and low iodine on biochemical indexes of the brain and learning-memory of offspring rats. *Fluoride* 37: 201-208.

Wang J, Zhang Y, Guo Z, Li R, Xue X, Sun Z, Niu R. 2018. Effects of perinatal fluoride exposure on the expressions of miR-124 and miR-132 in hippocampus of mouse pups. *Chemosphere* 197: 117-122.

Wei N, Dong Y, Wang Y, Guan Z. 2014. [Effects of chronic fluorosis on neurobehavioral development in offspring of rats and antagonistic effect of vitamin E]. *Chin J Endemiol* 33: 125-128.

Whitford GM, Whitford JL, Hobbs SH. 2009. Appetitive-based learning in rats: Lack of effect of chronic exposure to fluoride. *Neurotoxicol Teratol* 31: 210-215.

Wu CX, Gu XL, Ge YM, Zhang JH, Wang JD. 2006. Effects of high fluoride and arsenic on brain biochemical indexes and learning-memory in rats. *Fluoride* 39: 274-279.

Wu NP, Zhao ZL, Gao WH, Li XQ. 1995. [Behavioral teratology in rats exposed to fluoride.] *Chin J Endemiol* 12(5): 271-273.

Wu NP, Zhao ZL, Gao WH, Li XQ. 2008. Behavioral teratology in rats exposed to fluoride. *Fluoride* 41: 129-133.

Xu X, Shen X, Zhang Z. 2001. Effect of fluorosis on mice learning and memory behaviors and brain SOD activity and MDA content *China Public Health* 17(1): 8-10.

Yang L, Jin P, Wang X, Zhou Q, Lin X, Xi S. 2018. Fluoride activates microglia, secretes inflammatory factors and influences synaptic neuron plasticity in the hippocampus of rats. *Neurotox* 69: 108-120.

Yu Q, Shao D, Zhang R, Ouyang W, Zhang Z. 2019. Effects of drinking water fluorosis on L-type calcium channel of hippocampal neurons in mice. *Chemosphere* 220: 169-175.

Yuan J, Li Q, Niu R, Wang J. 2019. Fluoride exposure decreased learning ability and the expressions of the insulin receptor in male mouse hippocampus and olfactory bulb. *Chemosphere* 224: 71-76.

Zhang C, Ren C, Chen H, Geng R, Fan H, Zhao H, Guo K, Geng D. 2013. The analog of Ginkgo biloba extract 761 is a protective factor of cognitive impairment induced by chronic fluorosis. *Biol Trace Elem Res* 153: 229-236.

Zhang KL, Lou DD, Guan ZZ. 2015. Activation of the AGE/RAGE system in the brains of rats and in SH-SY5Y cells exposed to high level of fluoride might connect to oxidative stress. *Neurotoxicol Teratol* 48: 49-55.

Zhang J, Zhu W, Zhang Z. 2009. [The effect of fluorine exposure of pregnant rats on the learning and memory capabilities of baby rats]. *Chinese Journal of Public Health* 25(11): 1347-1348.

Zhang J, Zhu WJ, Xu XH, Zhang ZG. 2011. Effect of fluoride on calcium ion concentration and expression of nuclear transcription factor kappa-B rho65 in rat hippocampus. *Exp Toxicol Pathol* 63: 407-411.

Zhang J, Zhang Z. 2013. Effects of chronic fluorosis on camkiiα, c-FOS, BAX, and BCL-2 channel signaling in the hippocampus of rats. *Fluoride* 46: 135-141.

Zhang Z, Shen X, Xu X. 2001. [Effects of selenium on the damage of learning-memory ability of mice induced by fluoride]. *J Hyg Res* 30: 144-146.

Zhang Z, Xu X, Shen X, Xua XH. 1999. [Effect of fluoride exposure on synaptic structure of brain areas related to learning-memory in mice]. *J Hyg Res* 28(4): 210-212.

Zhang Z, Xu X, Shen X, Xua XH. 2008. Effect of fluoride exposure on synaptic structure of brain areas related to learning-memory in mice. *Fluoride* 41: 139-143.

Zhao Q, Niu Q, Chen J, Xia T, Zhou G, Li P, Dong L, Xu C, Tian Z, Luo C, Liu L, Zhang S, Wang A. 2019. Roles of mitochondrial fission inhibition in developmental fluoride neurotoxicity: Mechanisms of action in vitro and associations with cognition in rats and children. *Arch Toxicol* 93(3): 709-726.

Zhao XL, Wu JH. 1998. Actions of sodium fluoride on acetylcholinesterase activities in rats. *Biomed Environ Sci* 11: 1-6.

Zheng X, Sun Y, Ke L, Ouyang W, Zhang Z. 2016. Molecular mechanism of brain impairment caused by drinking-acquired fluorosis and selenium intervention. *Environ Toxicol Pharmacol* 43: 134-139.

Zhu W, Zhang J, Zhang Z. 2011. Effects of fluoride on synaptic membrane fluidity and PSD-95 expression level in rat hippocampus. *Biol Trace Elem Res* 139: 197-203.

Zhu YL, Zheng YJ, LV XM, Ma Y, Zhang J. 2012. Effects of fluoride exposure on performance in water labyrinth and monoamine neurotransmitters of rats. *Journal of Xinjiang Medical University* 3: 014.

Zhu YP, Xi SH, Li MY, Ding TT, Liu N, Cao FY, Zeng Y, Liu XJ, Tong JW, Jiang SF. 2017. Fluoride and arsenic exposure affects spatial memory and activates the ERK/CREB signaling pathway in offspring rats. *Neurotox* 59: 56-64.

Studies Not Available in HAWC

Abdelaleem MM, El-Tahawy NFG, Abozaid SMM, Abdel-Hakim SA. 2018. Possible protective effect of curcumin on the thyroid gland changes induced by sodium fluoride in albino rats: Light and electron microscopic study. *Endocr Regul* 52: 59-68.

Abd-Elhakim YM, Mohammed AT, Ali HA. 2018. Impact of subchronic exposure to triclosan and/or fluoride on estrogenic activity in immature female rats: The expression pattern of calbindin-D9k and estrogen receptor alpha genes. *J Biochem Mol Toxicol* 32(2): 22027.

Abdumajidov OR. 2004. [Sex differences in lipid peroxidation and antioxidant defense of the brain tissue in intoxication with low doses of inorganic compounds]. *Uzbekiston Tibbiet Zhurnali*: 58-60.

Adebayo OL, Shallie PD, Salau BA, Ajani EO, Adenuga GA. 2013. Comparative study on the influence of fluoride on lipid peroxidation and antioxidants levels in the different brain regions of well-fed and protein undernourished rats. *J Trace Elem Med Biol* 27: 370-374.

Adedara IA, Ojuade TJD, Olabiyi BF, Idris UF, Onibiyo EM, Ajeigbe OF, Farombi EO. 2016. Taurine ameliorates renal oxidative damage and thyroid dysfunction in rats chronically exposed to fluoride. *Biol Trace Elem Res*: 1-8.

Ahmed SK, Kalleny NK, Attia AAEM, Elkateb LA. 2015. The possible protective role of chromium chloride against sodium fluoride-induced changes in the structure of the cerebellar cortex of the adult male albino rat. *Egypt J Histol* 38: 402-414.

Al Badawi MH, Mahmoud OM, Salem NA. 2016. Therapeutic potential of omega-3 against sodium fluoride toxicity on the cerebellar cortex of adult male albino rats: Histological and immunohistochemical study. *Egypt J Histol* 39: 170-178.

Alhayani A, Elshal EB, Aal IHA, Al-Shammeri E, Kabra H. 2013. Does vitamin E protect against sodium fluoride toxicity on the cerebellar cortex of albino rats? *Middle East J Sci Res* 16: 1019-1026.

Ameeramja J, Raghunath A, Perumal E. 2018. Tamarind seed coat extract restores fluoride-induced hematological and biochemical alterations in rats. *Environ Sci Pollut Res Int* 25(26): 26157-26166.

Antonyan OA. 1980. [Lipid per oxidation in fluorosis and the protective role of dietary factors]. *Zh Eksp Klin Med* 20: 381-388.

Ardelean I, Racoveanu N, Manescu S, Lupulescu A, Diaconescu M, Ghelerter L. 1964. Experimental investigations concerning the effect of fluorine on the thyroid gland. *Rum Med Rev* 8: 17-20.

Atmaca N, Atmaca HT, Kanici A, Anteplioglu T. 2014. Protective effect of resveratrol on sodium fluoride-induced oxidative stress, hepatotoxicity and neurotoxicity in rats. *Food Chem Toxicol* 70: 191-197.

Auskaps AM, Shaw JH. 1955. Hemoglobin concentration, thyroid weight and growth rate in rats during minimum fluoride ingestion. *J Nutr* 55: 611-621.

Bagmut I, Kolisnyk I, Titkova A, Petrenko T, Filipchenko S. 2018. Content of catecholamines in blood serum of rats under fluoride intoxication. *Georgian Med News* (280-281): 125-129.

Bakalyan PH, Antonyan OA. 1981. [Effect of fluorosis on glutathione peroxidase and glutathione reductase activities and sulfhydryl groups]. *Zh Eksp Klin Med* 21: 10-14.

Basha PM, Madhusudhan N. 2010. Pre and post natal exposure of fluoride induced oxidative macromolecular alterations in developing central nervous system of rat and amelioration by antioxidants. *Neurochem Res* 35: 1017-1028.

Basha PM, Madhusudhan N. 2011. Effect of maternal exposure of fluoride on oxidative stress markers and amelioration by selected antioxidants in developing central nervous system of rats. *Biologia* 66: 187-193.

Basha PM, Rai P, Begum S. 2011. Evaluation of fluoride-induced oxidative stress in rat brain: A multigeneration study. *Biol Trace Elem Res* 142: 623-637.

Basha PM, Sujitha NS. 2012. Combined influence of intermittent exercise and temperature stress on the modulation of fluoride toxicity. *Biol Trace Elem Res* 148: 69-75.

Basha PM, Saumya SM. 2013. Suppression of mitochondrial oxidative phosphorylation and TCA enzymes in discrete brain regions of mice exposed to high fluoride: Amelioration by panax ginseng (ginseng) and lagerstroemia speciosa (banaba) extracts. *Cell Mol Neurobiol* 33: 453-464.

Basha MP, Begum S, Madhusudhan N. 2014. Antioxidants in the management of fluoride induced neural oxidative stress in developing rats. *Int J Pharm Sci Res* 5: 201-206.

Benetato G, Giuran AM, Cirmaciu R, Cirje M, Petrescu A, Vacariu A. 1970. [Effect of fluorine in drinking water on the metabolism of Ca and Mg and on neuromuscular excitability: Experimental studies and clinical observations]. *Rev Roum Physiol* 7: 335-352.

Bharti VK, Srivastava RS. 2009. Fluoride-induced oxidative stress in rat's brain and its amelioration by buffalo (Bubalus bubalis) pineal proteins and melatonin. *Biol Trace Elem Res* 130: 131-140.

Bhatnagar M, Rao P, Saxena A, Bhatnagar R, Meena P, Barbar S, Chouhan A, Vimal S. 2006. Biochemical changes in brain and other tissues of young adult female mice from fluoride in their drinking water. *Fluoride* 39: 280-284.

Bilgili A, Akdogan M, Yildiz M, Eraslan G, Cetin N. 2004. The effects of fluoride on thyroid hormones in rabbits. *Indian Vet J* 81: 986-988.

Bobek S, Kahl S, Ewy Z. 1976. Effect of long-term fluoride administration on thyroid hormones level blood in rats. *Endocrinol Exp* 10: 289-295.

Bouaziz H, Ammar E, Ghorbel H, Ketata S, Jamoussi K, Ayadi F, Guermazi F, Zeghal N. 2004. Effect of fluoride ingested by lactating mice on the thyroid function and bone maturation of their suckling pups. *Fluoride* 37: 133-142.

Bouaziz H, Soussia L, Guermazi F, Zeghal N. 2005. Fluoride-induced thyroid proliferative changes and their reversal in female mice and their pups. *Fluoride* 38: 185-192.

Bouaziz HB, Amara I, Essefi M, Croute F, Zeghal N. 2010. Fluoride-induced brain damages in suckling mice. *Pestic Biochem Physiol* 96: 24-29.

Chauhan SS, Ojha S, Mahmood A. 2013. Effects of fluoride and ethanol administration on lipid peroxidation systems in rat brain. *Indian J Exp Biol* 51: 249-255.

Chen J, Chen X, Yang K, Xia T, Xie H. 2002. [Studies on DNA damage and apoptosis in rat brain induced by fluoride]. *Chin J Prev Med* 36: 222-224.

Chirumari K, Reddy PK. 2007. Dose-dependent effects of fluoride on neurochemical milieu in the hippocampus and neocortex of rat brain. *Fluoride* 40: 101-110.

Chouhan S, Yadav A, Kushwah P, Kaul RK, Flora SJS. 2011. Silymarin and quercetin abrogates fluoride induced oxidative stress and toxic effects in rats. *Mol Cell Toxicol* 7: 25-32.

Clay AB, Suttie JW. 1987. Effect of dietary fluoride on dairy cattle: Growth of young heifers. *J Dairy Sci* 70: 1241-1251.

Czechowicz K, Osada A, Slesak B. 1974. Histochemical studies on the effect of sodium fluoride on metabolism in Purkinje's cells. *Folia Histochem Cytochem* 12: 37-44.

Demole V, Lerch P. 1956. [Normality of fixation of radioactive iodine in the thyroid of rats during experimental fluorosis]. *Helv Physiol Pharmacol Acta* 14(4): 62-63.

Dhurvey V, Patil V, Thakare M. 2017. Effect of sodium fluoride on the structure and function of the thyroid and ovary in albino rats (rattus norvegicus). *Fluoride* 50: 235-246.

Domzalska E. 1966. [Influence of sodium fluoride on hypophysis, thyroid gland, parathyroid, and adrenal gland in the white rat]. *Czas Stomatol* 19: 839-844.

El-lethey HS, Kamel MM, Shaheed IB. 2011. Perinatal exposure to sodium fluoride with emphasis on territorial aggression, sexual behaviour and fertility in male rats. *Life Sci J* 8: 686-694.

Flora SJS, Mittal M, Mishra D. 2009. Co-exposure to arsenic and fluoride on oxidative stress, glutathione linked enzymes, biogenic amines and DNA damage in mouse brain. *J Neurol Sci* 285: 198-205.

Flora SJS, Mittal M, Pachauri V, Dwivedi N. 2012. A possible mechanism for combined arsenic and fluoride induced cellular and DNA damage in mice. *Metallomics* 4: 78-90.

Gabovich RD, Verzhikovskaia NV. 1958. [Effect of fluoride compounds on absorption of radioactive iodine by the thyroid gland in humans and in experimental conditions]. *Probl Endokrinol Gormonoter* 4: 49-54.

Galamini-Ligori M, Di Blasi F. 1961. [Action of sodium fluoride on the thyroid of hypophysectomized rats]. *Boll Soc Ital Biol Sper* 37: 1503-1506.

Galletti P, Held HR, Korrodi H, Wegmann T. 1956. [Investigations on the possible damaging effect of fluorine on the thyroid gland]. *Helv Med Acta* 23: 601-605.

Ge Y, Ning H, Feng C, Wang H, Yan X, Wang S, Wang J. 2006. Apoptosis in brain cells of offspring rats exposed to high fluoride and low iodine. *Fluoride* 39: 173-178.

Ge Y, Niu R, Zhang J, Wang J. 2011. Proteomic analysis of brain proteins of rats exposed to high fluoride and low iodine. *Arch Toxicol* 85: 27-33.

Ge YM, Ning HM, Wang SL, Wang JD. 2005. DNA damage in thyroid gland cells of rats exposed to long-term intake of high fluoride and low iodine. *Fluoride* 38: 318-323.

Ge YM, Ning HM, Wang SL, Wang JD. 2005. Comet assay of DNA damage in brain cells of adult rats exposed to high fluoride and low iodine. *Fluoride* 38: 209-214.

Ge YM, Ning HM, Wang SL, Wang JD. 2005. Effects of high fluoride and low iodine on brain histopathology in offspring rats. *Fluoride* 38: 127-132.

Ge YM, Ning HM, Gu XL, Yin M, Yang XF, Qi YH, Wang JD. 2013. Effects of high fluoride and low iodine on thyroid function in offspring rats. *J Integr Agric* 12: 502-508.

Guan ZZ. 1986. [Morphology of the brain of the offspring of rats with chronic fluorosis]. *Chin J Pathol* 15: 297-299.

Guan Z, Wang Y, Xiao K. 1997. [Influence of experimental fluorosis on phospholipid content and fatty acid composition in rat brain]. *Chin Med J* 77: 592-596.

Guan Z-Z, Wang Y-N, Xiao K-Q, Dai D-Y, Chen Y-H, Liu J-L, Sindelar P, Dallner G. 1998. Influence of chronic fluorosis on membrane lipids in rat brain. *Neurotoxicol Teratol* 20: 537-542.

Guan ZZ, Shan KR, Wang YN, Dallner G. 2006. [Changes of lipids and nicotinic receptors in rat brains and pheochromocytoma with fluorosis]. *Chin J Endemiol* 25: 121-124.

Gushchin SK. 1951. [Effect of sodium fluoride on iodine metabolism in rabbit tissue organs; on the etiology of endemic goiter]. *Gig Sanit* 2: 45-48.

Hamza RZ, Al-Harbi MS. 2014. Sodium fluoride induced neurotoxicity and possible antioxidant role of selenium and curcumin in male mice. *Biosci Biotechnol Res Asia* 11: 81-87.

Hamza RZ, El-Shenawy NS, Ismail HAA. 2015. Protective effects of blackberry and quercetin on sodium fluoride-induced oxidative stress and histological changes in the hepatic, renal, testis and brain tissue of male rat. *J Basic Clin Physiol Pharmacol* 26: 237-251.

Haojun Z, Yaoling W, Ke Z, Jin L, Junling W. 2012. Effects of NaF on the expression of intracellular Ca2+ fluxes and apoptosis and the antagonism of taurine in murine neuron. *Toxicol Mech Methods* 22: 305-308.

Hara K. 1980. Studies on fluorosis especially effects of fluoride on thyroid metabolism. *J Dent Health* 30: 42-57.

Harris NO, Hayes RL. 1955. A tracer study of the effect of acute and chronic exposure to sodium fluoride on the thyroid iodine metabolism of rats. *J Dent Res* 34: 470-477.

Hassan HA, Abdel-Aziz AF. 2010. Evaluation of free radical-scavenging and anti-oxidant properties of black berry against fluoride toxicity in rats. *Food Chem Toxicol* 48: 1999-2004.

Hoogstratten B, Leone NCLG, Shupe J, Greenwood DA, Lieberman J. 1965. Effect of fluorides on hematopoietic system, liver, and thyroid gland in cattle. *J Amer Med Assoc* 192: 26-32.

Inkielewicz I, Rogowska M, Krechniak J. 2006. Lipid peroxidation and antioxidant enzyme activity in rats exposed to fluoride and ethanol. *Fluoride* 39: 53-59.

Inkielewicz I, Czarnowski W. 2008. Oxidative stress parameters in rats exposed to fluoride and aspirin. *Fluoride* 41: 76-82.

Inkielewicz-Stepniak I, Czarnowski W. 2010. Oxidative stress parameters in rats exposed to fluoride and caffeine. *Food Chem Toxicol* 48: 1607-1611.

Jiang P, Li G, Zhou X, Wang C, Qiao Y, Liao D, Shi D. 2018. Chronic fluoride exposure induces neuronal apoptosis and impairs neurogenesis and synaptic plasticity: Role of GSK-3beta/beta-catenin pathway. *Chemosphere* 214: 430-435.

Jiang SF, Xi SH, Yao SQ, Tong JW, Zhang YS, Wang Q, Su J, Li MY. 2013. [Effects of fluoride, arsenic and co-exposure on expression of Bcl-2 and Bax in hippocampus and cerebral cortex of rats]. *Chin J Endemiol* 32: 365-369.

Jiang Y, Guo X, Sun Q, Shan Z, Teng W. 2016. Effects of excess fluoride and iodide on thyroid function and morphology. *Biol Trace Elem Res* 170: 382-389.

Jin TX, Guan ZZ, Zhang H. 2011. [The effect of fluoride on α subunit of calcium/calmodulin-dependent protein kinase-II mRNA and protein expression in central nervous system]. *Chin J Endemiol* 30: 247-250.

Jonderko G, Kita K, Pietrzak J, Primus-Slowinska B, Ruranska B, Zylka-Wloszczyk M, Straszecka J. 1983. [Effect of subchronic sodium fluoride poisoning on the thyroid gland of rabbits with normal and increased supply of iodine]. *Endokrynol Pol* 34: 195-203.

Kahl S, Bobek S. 1975. [Effect of fluoride administration on radiothyroxine turnower in rats]. *Endokrynol Pol* 26: 391-396.

Kahl S, Ewy Z. 1975. Effect of single and long term sodium fluoride administration on biosynthesis of the thyroid hormone in rats. *Fluoride* 8: 191-198.

Kapoor V, Prasad T, Paliwal VK. 2001. Blood biochemical constituents in calves following subclinical levels of fluoride toxicosis. *Fluoride* 34: 126-131.

Karawya FS, Zahran NM, Azzam EZ. 2015. Is water fluoridation a hidden cause of obesity? Histological study on thyroid follicular cells of albino rats. *Egypt J Histol* 38: 547-557.

Kaur T, Bijarnia RK, Nehru B. 2009. Effect of concurrent chronic exposure of fluoride and aluminum on rat brain. *Drug Chem Toxicol* 32: 215-221.

Kelimu A, Liu KT, Lian J, Hu HH, Zheng YJ, Wang TM. 2008. [Effects of vitamin C and E on the ultrastructure in liver, kidney and brain of fluorosis rats]. *Chin J Endemiol* 27: 378-381.

Kinawy AA. 2019. Synergistic oxidative impact of aluminum chloride and sodium fluoride exposure during early stages of brain development in the rat. *Environ Sci Pollut Res Int* 26(11): 10951-10960.

Knizhnikov VA. 1959. [Effect of potable water with high fluoride concentration on thyroid function]. *Gig Sanit* 24: 20-25.

Knizhnikov VA, Tsypin AB, Shcherbova EN, Bugryshev PF. 1963. [The effect of drinking water with an increased fluorine content on the bioelectrical activity of the brain and heart under experimental conditions]. *Gig Sanit* 28: 16-19.

Kondo T, Yoshida M, Kasahara K. 1976. [Acute fluorosis in female rats: Time of inhibition and recovery of cholinesterase in serum and salivary glands]. *Jpn J Dent Health* 26: 187-192.

Kowalewska M. 1974. [Biopotentials of the organ of hearing in chronic sodium fluoride poisoning]. *J Pol Otolaryngol* 28: 417-424.

Krechniak J, Inkielewicz I. 2005. Correlations between fluoride concentrations and free radical parameters in soft tissues of rats. *Fluoride* 38: 293-296.

Leonard BE. 1972. Effect of phentolamine on the increase in brain glycolysis following the intraventricular administration of dibutyryl-3,5-cyclic adenosine monophosphate and sodium fluoride to mice. *Biochem Pharmacol* 21: 115-117.

Liu G, Zhang W, Jiang P, Li X, Liu C, Chai C. 2012. Role of nitric oxide and vascular endothelial growth factor in fluoride-induced goitrogenesis in rats. *Environ Toxicol Pharmacol* 34: 209-217.

Li H, Cai Q, Wang D. 2012. [Effect of fluoride on the expression of rat thyroid peroxidase mRNA]. *Chin J Endemiol* 31: 515-517.

Li H, Cai Q, Wang D. 2012. [Effects of fluoride on rat thyroid morphology, thyroid peroxidase activity and the expression of thyroid peroxidase protein]. *Chin J Endemiol* 31: 271-274.

Liu H, Hou C, Zeng Q, Zhao L, Cui Y, Yu L, Wang L, Zhao Y, Nie J, Zhang B, Wang A. 2016. Role of endoplasmic reticulum stress-induced apoptosis in rat thyroid toxicity caused by excess fluoride and/or iodide. *Environ Toxicol Pharmacol* 46: 277-285.

Liu YJ, Gao Q, Wu CX, Guan ZZ. 2010. [Changes of the c-Jun N-terminal kinase in the brains of rats with chronic fluorosis]. *Chin J Endemiol* 29: 608-612.

Liu YJ, Guan ZZ, Gao Q, Pei JJ. 2011. Increased level of apoptosis in rat brains and SH-SY5Y cells exposed to excessive fluoride-A mechanism connected with activating JNK phosphorylation. *Toxicol Lett* 204: 183-189.

Lohakare J, Pattanaik AK. 2013. Effects of addition of fluorine in diets differing in protein content on urinary fluoride excretion, clinical chemistry and thyroid hormones in calves. *Brazilian J Anim Sci* 42: 751-758.

Long YG, Wang YN, Chen J, Jiang SF, Nordberg A, Guan ZZ. 2002. Chronic fluoride toxicity decreases the number of nicotinic acetylcholine receptors in rat brain. *Neurotoxicol Teratol* 24: 751-757.

Lou DD, Liu YF, Zhang KL, Yu YN, Guan ZZ. 2011. [Changes of reactive oxygen species level and mitochondria fission-fusion in cortical neurons of rats with chronic fluorosis]. *Chin J Endemiol* 30: 256-260.

Lou DD, Liu YF, Qin SL, Zhang KL, Yu YN, Guan ZZ. 2012. [Changed transcription level of mitochondrial fission and fusion gene loci in cortical neurons of rats with chronic fluorosis]. *Chin J Endemiol* 31: 125-129.

Lou DD, Zhang KL, Qin SL, Liu YF, Yu YN, Guan ZZ. 2012. [Alteration of mitochondrial distribution and gene expression of fission 1 protein in cortical neurons of rats with chronic fluorosis]. *Chin J Pathol* 41: 243-247.

Lou DD, Pan JG, Zhang KL, Qin SL, Liu YF, Yu YN, Guan ZZ. 2013. [Changed expression of mito-fusion 1 and mitochondrial fragmentation in the cortical neurons of rats with chronic fluorosis]. *Chin J Prev Med* 47: 170-174.

Lou DD, Zhang KL, Pan JG, Qin SL, Liu YF, Yu YN, Guan ZZ. 2013. [Influence of chronic fluorosis on the expression of mitochondrial fission protein dynamin-related 1 in the cortical neurons of rats]. *Chin J Prev Med* 47: 561-564.

Lou DD, Zhang KL, Qin SL, Liu YF, Liu YJ, Guan ZZ. 2013. [Effects of chronic fluorosis on 4.8 kb mitochondrial DNA in liver, kidney and brain of rats]. *Chin J Endemiol* 32: 121-124.

Lou DD, Guan ZZ, Pei JJ. 2014. Alterations of apoptosis and expressions of Bax and Bcl-2 in the cerebral cortices of rats with chronic fluorosis. *Fluoride* 47: 199-207.

Luo GY, Niu RY, Sun ZL, Zhang JH, Wang JM, Wang C, Wang JD. 2011. Reduction of CaMKII expression in the hippocampus of rats from ingestion of fluoride and/or lead. *Fluoride* 44: 63-69.

Ma T, Liu D, Song K. 1999. Cytochemical study of neuron enzyme at anterior horn of spinal cord in rats with experimental fluorosis. *J Chin Med Univ* 28: 81-82.

Ma TX, Yu HT, Song KQ. 2008. [Expression of c-fos and Caspase 8 in cerebral cortex of rats with experimental fluorosis]. *Chin J Endemiol* 27: 131-133.

Mach Z, Zygulska-Machowa H. 1959. O wplywie fluoru na przemiane J131 [Russian and English summ.]. *Endokrynol Pol* 10: 157-162.

Machida H. 1989. [A study on the rabbit thermoregulatory system effects of high dose sodium fluoride]. *Dent Sci Rep* 89: 607-626.

Madan J, Puri JP, Singh JK. 2009. Growth, feed efficiency and blood profile of buffalo calves consuming high levels of fluoride. *Trop Anim Health Prod* 41: 295-298.

Madhusudhan N, Basha PM, Begum S, Ahmed F. 2009. Fluoride-induced neuronal oxidative stress and its amelioration by antioxidants in developing rats. *Fluoride* 42: 179-187.

Madhusudhan N, Basha PM, Rai P, Ahmed F, Prasad GR. 2010. Effect of maternal fluoride exposure on developing CNS of rats: Protective role of Aloe vera, Curcuma longa and Ocimum sanctum. *Indian J Exp Biol* 48: 830-836.

Manocha SL, Warner H, Olkowski ZL. 1975. Cytochemical response of kidney, liver and nervous system of fluoride ions in drinking water. *Histochem J* 7: 343-355.

Mansour HH, Tawfik SS. 2012. Efficacy of lycopene against fluoride toxicity in rats. *Pharm Biol* 50: 707-711.

Mietkiewski K, Walczak M, Trojanowicz R. 1966. [Effect of sodium fluoride on the neurosecretory system in guinea pigs]. *Endokrynol Pol* 17: 121-131.

Mohamed NE. 2016. The role of calcium in ameliorating the oxidative stress of fluoride in rats. *Biol Trace Elem Res* 170: 128-144.

Muhlemann HR, Schneider R. 1956. [Mitotic activity of rat thyroid epithelium after administration of fluoridated drinking water]. *Schweiz Med Wochenschr* 86: 625-627.

Nabavi SF, Eslami S, Moghaddam AH, Nabavi SM. 2011. Protective effects of curcumin against fluoride-induced oxidative stress in the rat brain. *Neurophysiology* 43: 287-291.

Nabavi SF, Moghaddam AH, Nabavi SM, Eslami S. 2011. Protective effect of curcumin and quercetin on thyroid function in sodium fluoride intoxicated rats. *Fluoride* 44: 147-152.

Nabavi SF, Habtemariam S, Jafari M, Sureda A, Nabavi SM. 2012. Protective role of gallic acid on sodium fluoride induced oxidative stress in rat brain. *Bull Environ Contam Toxicol* 89: 73-77.

Nabavi SF, Nabavi SM, Latifi AM, Mirzaei M, Habtemariam S, Moghaddam AH. 2012. Mitigating role of quercetin against sodium fluoride-induced oxidative stress in the rat brain. *Pharm Biol* 50: 1380-1383.

Nabavi SF, Nabavi SM, Habtemariam S, Moghaddam AH, Sureda A, Mirzaei M. 2013. Neuroprotective effects of methyl-3-O-methyl gallate against sodium fluoride-induced oxidative stress in the brain of rats. *Cell Mol Neurobiol* 33: 261-267.

Nabavi SM, Sureda A, Nabavi SF, Latifi AM, Moghaddam AH, Hellio C. 2012. Neuroprotective effects of silymarin on sodium fluoride-induced oxidative stress. *J Fluor Chem* 142: 79-82.

Narayanaswamy M, Piler MB. 2010. Effect of maternal exposure of fluoride on biometals and oxidative stress parameters in developing CNS of rat. *Biol Trace Elem Res* 133: 71-82.

Narbutt B, Romer TE, Grabski J, Szymik N. 1971. [Influence of natrium fluoride on the structure of the rat thyroid]. *Endokrynol Pol* 22: 445-451.

Narbutt B, Romer TE, Grabski J, Szymik N. 1971. [The influence of sodium fluoride on the morphology of the thyroid gland in rats]. *Endokrynol Pol* 22: 361-365.

Niu RY, Sun ZL, Cheng ZT, Liu HT, Chen HC, Wang JD. 2008. Effects of fluoride and lead on N-methyl-D-aspartate receptor 1 expression in the hippocampus of offspring rat pups. *Fluoride* 41: 101-110.

Niu R, Wang J, Sun Z, Xue X, Yan X, Zhang J. 2015. Transcriptional regulatory dynamics of the hypothalamic-pituitary-testicular axis in male mice exposed to fluoride. *Environ Toxicol Pharmacol* 40: 557-562.

Niu R, Zhang Y, Liu S, Liu F, Sun Z, Wang J. 2015. Proteome alterations in cortex of mice exposed to fluoride and lead. *Biol Trace Elem Res* 164: 99-105.

Ogilvie AL. 1952. Histological findings in the kidney, liver, pancreas, adrenal and thyroid gland of the rat following sodium fluoride administration. *J Dent Res* 31: 598-598.

Okayasu I, Tsuchida M, Yanagisawa F. 1985. Hyperplastic nodules of thyroid parafollicular cells (C cells) in rats induced by prolonged low dose ingestion of NaF. *Fluoride* 18: 111-117.

Pal S, Sarkar C. 2014. Protective effect of resveratrol on fluoride induced alteration in protein and nucleic acid metabolism, DNA damage and biogenic amines in rat brain. *Environ Toxicol Pharmacol* 38: 684-699.

Pan Y, Lu P, Yin L, Chen K, He Y. 2015. Effect of fluoride on the proteomic profile of the hippocampus in rats. *Z Naturforsch C* 70: 151-157.

Phillips PH, Lamb AR. 1934. Histology of certain organs and teeth in chronic toxicosis due to fluorin. *Arch Path* 17: 169-176.

Portela ML. 1972. [Biochemical effects in the prolonged ingestion of fluorides in rats]. *Arch Latinoam Nutr* 22: 291-308.

Prestes DS, Filappi A, Schossler DR, Duarte FA, Dressler VL, Flores EMM, Cecim M. 2009. Functional and histological evaluations of thyroid of sheep submitted to sodium fluoride administration. *Arq Bras Med Vet Zootec* 61: 293-298.

Puentes F, Cremer HD. 1966. Experiments on fluoride-iodine antagonism in the thyroid gland. *Adv Fluorine Res* 4: 213-220.

Qian W, Miao K, Li T, Zhang Z. 2013. Effect of selenium on fluoride-induced changes in synaptic plasticity in rat hippocampus. *Biol Trace Elem Res* 155: 253-260.

Qing-Feng S, Ying-Peng X, Tian-Tong X. 2019. Matrix metalloproteinase-9 and p53 involved in chronic fluorosis induced blood-brain barrier damage and neurocyte changes. *Arch Med Sci* 15(2): 457-466.

Qiu YH, Kong DM, Yang Q, Zhao N. 2010. [Influence of high-fluoride on thyroid function and brain damage in rats]. *Chin J Endemiol* 29: 146-149.

Raghavendra M, Ravindra RK, Raghuveer YP, Narasimha JK, Uma MRV, Navakishor P. 2016. Alleviatory effects of hydroalcoholic extract of cauliflower (brassica oleracea var. botrytis) on thyroid function in fluoride intoxicated rats. *Fluoride* 49: 84-90.

Rakhov GM. 1964. [Effect of calcium and fluorine in drinking water on the iodine metabolism and status of the thyroid gland in iodine insufficiency in food]. *Gig Sanit* 29: 12-17.

Ranpariya VL, Parmar SK, Sheth NR, Chandrashekhar VM. 2011. Neuroprotective activity of matricaria recutita against fluoride-induced stress in rats. *Pharm Biol* 49: 696-701.

Reddy KP, Sailaja G, Krishnaiah C. 2009. Protective effects of selenium on fluoride induced alterations in certain enzymes in brain of mice. *J Environ Biol* 30: 859-864.

Rogalska A, Kuter K, Zelazko A, Glogowska-Gruszka A, Swietochowska E, Nowak P. 2017. Fluoride alteration of [3H]glucose uptake in Wistar rat brain and peripheral tissues. *Neurotoxicol Res* 31: 436-443.

Saka O, Hallac P, Urgancioğlu I. 1965. The effect of fluoride on the thyroid of the rat. *New Istanbul Contrib Clin Sci* 8: 87-90.

Samanta A, Chanda S, Bandyopadhyay B, Das N. 2016. Establishment of drug delivery system nanocapsulated with an antioxidant (+)-catechin hydrate and sodium meta borate chelator against sodium fluoride induced oxidative stress in rats. *J Trace Elem Med Biol* 33: 54-67.

Sarkar C, Das N, Pal S, Dinda B. 2014. Oxidative stress induced alteration of protein and nucleic acid metabolism in fluoride-intoxicated rat brain: Protection by 3α-hydroxy olean-12-en-27-oic acid isolated from neanotis wightiana. *Int J Pharm Sci Res* 5: 3047-3066.

Sarkar C, Pal S. 2014. Ameliorative effect of resveratrol against fluoride-induced alteration of thyroid function in male Wistar rats. *Biol Trace Elem Res* 162: 278-287.

Sarkar C, Pal S, Das N, Dinda B. 2014. Ameliorative effects of oleanolic acid on fluoride induced metabolic and oxidative dysfunctions in rat brain: Experimental and biochemical studies. *Food Chem Toxicol* 66: 224-236.

Seffner W, Teubener W, Runde H, Wiedner H, Vogt J, Otto G, Zschau E, Geinitz D, Franke J. 1990. Boron as an antidote to fluorosis? II. Studies on various organs of pigs. *Fluoride* 23: 68-79.

Selim AOA, El-Haleem MR, Ibrahim IH. 2012. Effect of sodium fluoride on the thyroid gland of growing male albino rats: Histological and biochemical study. *Egypt J Histol* 35: 470-482.

Shao Q, Wang Yn, Guan Z. 2000. [Influence of free radical inducer on the level of oxidative stress in brain of rats with fluorosis]. *Chin J Prev Med* 34: 330-332.

Sharma C, Suhalka P, Sukhwal P, Jaiswal N, Bhatnagar M. 2014. Curcumin attenuates neurotoxicity induced by fluoride: An in vivo evidence. *Pharmacogn Mag* 10: 61-65.

Shashi A. 1992. Studies on alterations in brain lipid metabolism following experimental fluorosis. *Fluoride* 25: 77-84.

Shashi A. 1993. Nucleic acid levels in thyroid gland in acute and chronic fluoride intoxication. *Fluoride* 26: 191-196.

Shashi A, Singh JP, Thapar SP. 1994. Effect of long-term administration of fluoride on levels of protein, free amino acids and RNA in rabbit brain. *Fluoride* 27: 155-159.

Shashi A. 2003. Histopathological investigation of fluoride-induced neurotoxicity in rabbits. *Fluoride* 36: 95-105.

Shashi A, Neetika S, Bhardwaj M. 2009. Neuronal DNA damage and apoptosis in brain of rat exposed to fluoride. *Asian J Microbiol Biotechnol Environ Sci* 11: 629-632.

Shayiq RM, Raza H, Kidwai AM. 1986. Fluoride and lipid peroxidation a comparative study in different rat tissues. *Bull Environ Contam Toxicol* 37: 70-76.

Shen QF, Li HN, Xu TT, Xia YP. 2012. [Damage of blood brain barrier of spinal cord in rats with chronic fluorosis]. *Chin Med J* 92: 2357-2361.

Shen Q, Tian R, Li H, Xu T, Xia Y. 2014. [White matter injury of spinal cord in rats with chronic fluorosis and recovery after defluoridation]. *Chin Med J* 94: 1189-1192.

Shen X, Zhang Z, Xu X. 2004. [Influence of combined iodine and fluoride on phospholipid and fatty acid composition in brain cells of rats]. *J Hyg Res* 33: 158-161.

Shivarajashankara YM, Shivashankara AR, Bhat PG, Rao SH. 2001. Effect of fluoride intoxication on lipid peroxidation and antioxidant systems in rats. *Fluoride* 34: 108-113.

Shivarajashankara YM, Shivashankara AR, Bhat PG, Rao SH. 2002. Brain lipid peroxidation and antioxidant systems of young rats in chronic fluoride intoxication. *Fluoride* 35: 197-203.

Shivarajashankara YM, Shivashankara AR, Bhat PG, Rao SM, Rao SH. 2002. Histological changes in the brain of young fluoride-intoxicated rats. *Fluoride* 35: 12-21.

Siebenhuner L, Miloni E, Burgi H. 1984. [Effects of fluoride on thyroid hormone biosynthesis: Studies in a highly sensitive test system]. *Klin Wochenschr* 62: 859-861.

Singh R, Srivastava AK, Gangwar NK. 2017. Clinico-pathological studies on the co-exposure of cypermethrin and fluoride in experimental rats with ameliorative action of Vitamin E. *Vet Pract* 18(2): 207-210.

Soni KK, Shrivastava VK. 2007. Sodium fluoride induced histopathological changes in thyroid gland of male mus musculus. *Biochem Cell Arch* 7: 317-320.

Stee EW. 1968. *Effect of sodium fluoride and AMOX (NF3O) on growth and thyroid function in the rat*. No. AMRL-TR-67-189. Wright-Patterson Air Force Base, OH: pp. 67.

Štolc V, Podoba J. 1960. Effect of fluoride on the biogenesis of thyroid hormones. *Nature* 188: 855-856.

Sugivama Y. 1967. [The effect of sodium fluoride administration on the parathyroid glands]. *Hirosaki Med J* 19: 520-529.

Sun Y, Ke L, Zheng X, Li T, Ouyang W, Zhang Z. 2016. Effects of different levels of calcium intake on brain cell apoptosis in fluorosis rat offspring and its molecular mechanism. *Biol Trace Elem Res*: 1-12.

Takata H. 1958. The effect of fluorine upon the uptake of I131 by the thyroid glands. *Folia Pharmacol Jpn* 54: 230-236.

Teng Y, Zhang J, Zhang Z, Feng J. 2017. The effect of chronic fluorosis on calcium ions and CaMKIIα, and c-fos expression in the rat hippocampus. *Biol Trace Elem Res*: 295-302.

Trabelsi M, Guermazi F, Zeghal N. 2001. Effect of fluoride on thyroid function and cerebellar development in mice. *Fluoride* 34: 165-173.

Tsuchida M, Okayasu I, Kohyama Y, Kurihara H, Tanaka H, Yanagisawa F, Date C, Hayashi M, Mui K, Asada M. 1986. Effects of long term, low dose ingestion of fluoride on the thyroid gland in rats. *Stud Environ Sci* 27: 307-312.

Vani ML, Reddy KP. 2000. Effects of fluoride accumulation on some enzymes of brain and gastrocnemius muscle of mice. *Fluoride* 33: 17-26.

Wang C, Liang C, Ma J, Manthari RK, Niu R, Wang J, Wang J, Zhang J. 2018. Co-exposure to fluoride and sulfur dioxide on histological alteration and DNA damage in rat brain. *J Biochem Mol Toxicol* 32.

Wang H, Yang Z, Zhou B, Gao H, Yan X, Wang J. 2009. Fluoride-induced thyroid dysfunction in rats: Roles of dietary protein and calcium level. *Toxicol Ind Health* 25: 49-57.

Wang J, Niu R, Sun Z, Lv L, Smith GW. 2008. Effects of protein and calcium supplementation on bone metabolism and thyroid function in protein and calcium deficient rabbits exposed to fluoride. *Fluoride* 41: 283-291.

Wang JD, Ge YM, Ning HM, Wang SL. 2004. Effects of high fluoride and low iodine on oxidative stress and antioxidant defense of the brain in offspring rats. *Fluoride* 37: 264-270.

Wang JL. 2007. [Effect of fluoride on the intracellular Ca2+ in neurons of mice]. *Chin J Endemiol* 26: 505-507.

Wang Y, Guan Z, Xiao K. 1997. [Changes of coenzyme Q content in brain tissues of rats with fluorosis]. *Chin J Prev Med* 31: 330-333.

Wang Y, Dong Y, Wei N, Guan Z. 2015. [Influence of chronic fluorosis on expression of quinone oxidoreductase-1 and heme oxygenase-1 in rat brains]. *Chin J Endemiol* 34: 250-253.

Wedzisz A, Cieciura J. 1988. Effect of small sodium fluoride feed supplements on the serum thyroid hormone content of rats. *Bromatol Chem Toksykol* 21: 174-175.

Wei N, Dong YT, Deng J, Wang Y, Qi XL, Yu WF, Xiao Y, Zhou JJ, Guan ZZ. 2018. Changed expressions of N-methyl-d-aspartate receptors in the brains of rats and primary neurons exposed to high level of fluoride. *J Trace Elem Med Biol* 45: 31-40.

Yan N, Liu Y, Liu S, Cao S, Wang F, Wang Z, Xi S. 2016. Fluoride-induced neuron apoptosis and expressions of inflammatory factors by activating microglia in rat brain. *Mol Neurobiol* 53: 4449-4460.

Yang H, Xing R, Liu S, Yu H, Li P. 2016. Gamma-Aminobutyric acid ameliorates fluoride-induced hypothyroidism in male Kunming mice. *Life Sci* 146: 1-7.

Yang H, Xing R, Liu S, Yu H, Li P. 2019. Analysis of the protective effects of gamma-aminobutyric acid during fluoride-induced hypothyroidism in male Kunming mice. *Pharm Biol* 57(1): 29-37.

Yang M, Ren Z, Zhou B, Guan Z, Yu W. 2017. [Expression of endonuclease G in the brain tissue of rats with chronic fluorosis]. *Chin J Endemiol* 36: 327-332.

Yuan SD, Xie QW, Lu FY. 1993. Changes of serotonin content and turnover rate in hypothalamus of female rat during fluorosis. *Fluoride* 26: 57-60.

Zhai JX, Guo ZY, Hu CL, Wang QN, Zhu QX. 2003. [Studies on fluoride concentration and cholinesterase activity in rat hippocampus]. *Chin J Ind Hyg Occup Dis* 21: 102-104.

Zhan CW, Huo DJ. 1988. Ultrastructural findings in liver, kidneys, thyroid-gland and cardiac-muscle of rabbits following sodium-fluoride administration. *Fluoride* 21: 32-38.

Zhan XA, Xu ZR, Li JX, Wang M. 2005. Effects of fluorosis on lipid peroxidation and antioxidant systems in young pigs. *Fluoride* 38: 157-161.

Zhan XA, Li JX, Wang M, Xu ZR. 2006. Effects of fluoride on growth and thyroid function in young pigs. *Fluoride* 39: 95-100.

Zhang KL, Lou DD, Liu YF, Qin SL, Guan ZZ. 2012. [Changes of P-glycoprotein and nuclear factor κB in the cerebral cortex of rat with chronic fluorosis]. *Chin J Endemiol* 31: 613-616.

Zhang KL, Lou DD, Guan ZZ. 2013. [Expression of receptor for advanced glycation endproducts and nuclear factor κB in brain hippocampus of rat with chronic fluorosis]. *Chin J Endemiol* 32: 625-628.

Zhang WD, Zhang Y, Liu GY, Jiang P, Chai CY. 2008. [Effects of fluoride on ultrastructure of thyroids in rats]. *Chin J Endemiol* 27: 622-624.

Zhang ZG, Wang XY, Nian WW, Liao QX, Zhang R, Ouyang W. 2017. Effects of calcium on drinking fluorosis-induced hippocampal syntaptic plasticity impairment in the offspring of rats. *Transl Neurosci* 8: 191-200.

Zhao W, Zhu H, Yu Z, Aoki K, Misumi J, Zhang X. 1998. Long-term effects of various iodine and fluorine doses on the thyroid and fluorosis in mice. *Endocr Regul* 32: 63-70.

Zhao WY. 1988. [A preliminary study of the interaction of iodine and fluoride in experimental iodine goiter and fluorosis]. *Chin J Prev Med* 22: 146-148.

Zhao XL, Gao WH, Zhao ZL. 1994. [Effects of sodium fluoride on the activity of Ca2+Mg(2+)-ATPase in synaptic membrane in rat brain]. *Chin J Prev Med* 28: 264-266.

Zhavoronkov AA, Polyakova GA. 1973. Morphological and functional state of the hypothalamo-hypophyseal neurosecretory system in experimental fluorosis. *Bull Exp Biol Med* 75: 194-196.

Zhou B, Luo G, Wang C, Niu R, Wang J. 2014. Effects of fluoride on expression of cytokines in the hippocampus of adult rats. *Fluoride* 47: 191-198.

In Vitro Experimental Studies

As described in **Figure 4**, 60 in vitro experimental studies were included; however, data extraction was not conducted on in vitro studies. Therefore, in vitro experimental studies are not available in HAWC with the exception of in vitro studies that also reported in vivo non-human animal data that meet the relevant criteria for being made available in HAWC. The following lists of references are organized as studies that are available in HAWC (n = 6) followed by studies that are not available in HAWC (n = 54).

Studies Available in HAWC

Chen J, Niu Q, Xia T, Zhou G, Li P, Zhao Q, Xu C, Dong L, Zhang S, Wang A. 2018. ERK1/2-mediated disruption of BDNF-TrkB signaling causes synaptic impairment contributing to fluoride-induced developmental neurotoxicity. *Toxicology* 410: 222-230.

Niu Q, Chen J, Xia T, Li P, Zhou G, Xu C, Zhao Q, Dong L, Zhang S, Wang A. 2018. Excessive ER stress and the resulting autophagic flux dysfunction contribute to fluoride-induced neurotoxicity. *Environ Pollut* 233: 889-899.

Shan KR, Qi XL, Long YG, Nordberg A, Guan ZZ. 2004. Decreased nicotinic receptors in PC12 cells and rat brains influenced by fluoride toxicity: A mechanism relating to a damage at the level in post-transcription of the receptor genes. *Toxicology* 200: 169-177.

Zhang KL, Lou DD, Guan ZZ. 2015. Activation of the AGE/RAGE system in the brains of rats and in SH-SY5Y cells exposed to high level of fluoride might connect to oxidative stress. *Neurotoxicol Teratol* 48: 49-55.

Zhao Q, Niu Q, Chen J, Xia T, Zhou G, Li P, Dong L, Xu C, Tian Z, Luo C, Liu L, Zhang S, Wang A. 2019. Roles of mitochondrial fission inhibition in developmental fluoride neurotoxicity: Mechanisms of action in vitro and associations with cognition in rats and children. *Arch Toxicol* 93(3): 709-726.

Zhao XL, Wu JH. 1998. Actions of sodium fluoride on acetylcholinesterase activities in rats. *Biomed Environ Sci* 11: 1-6.

Studies Not Available in HAWC

Ardelean I, Racoveanu N, Manescu S, Lupulescu A, Diaconescu M, Ghelerter L. 1964. Experimental investigations concerning the effect of fluorine on the thyroid gland. *Rum Med Rev* 8: 17-20.

Chen J, Chen X, Yang K. 2000. [Effects of selenium and zinc on the DNA damage caused by fluoride in pallium neural cells of rats]. *J Hyg Res* 29: 216-217.

Chen L, Ning H, Yin Z, Song X, Feng Y, Qin H, Li Y, Wang J, Ge Y, Wang W. 2017. The effects of fluoride on neuronal function occurs via cytoskeleton damage and decreased signal transmission. *Chemosphere* 185: 589-594.

Chen R, Zhao LD, Liu H, Li HH, Ren C, Zhang P, Guo KT, Zhang HX, Geng DQ, Zhang CY. 2017. Fluoride induces neuroinflammation and alters Wnt signaling pathway in BV2 microglial cells. *Inflammation* 40: 1123-1130.

Cheng TJ, Chen TM, Chen CH, Lai YK. 1998. Induction of stress response and differential expression of 70 kDa stress proteins by sodium fluoride in HeLa and rat brain tumor 9L cells. *J Cell Biochem* 69: 221-231.

Deng MF, Zhu D, Liu YP, He WW, Gui CZ, Guan ZZ. 2018. Attenuation by 7-nitroindazole of fluoride-induced toxicity in SH-SY5Y cells exposed to high fluoride: Effects on nitric oxide, nitric oxide synthetase activity, nNOS, and apoptosis. *Fluoride* 51(4): 328-339.

Flores-Mendez M, Ramirez D, Alamillo N, Hernandez-Kelly LC, Del Razo LM, Ortega A. 2014. Fluoride exposure regulates the elongation phase of protein synthesis in cultured Bergmann glia cells. *Toxicol Lett* 229: 126-133.

Gao Q, Liu YH, Guan ZZ. 2008. Oxidative stress might be a mechanism connected with the decreased alpha 7 nicotinic receptor influenced by high-concentration of fluoride in SH-SY5Y neuroblastoma cells. *Toxicol In Vitro* 22: 837-843.

Goschorska M, Gutowska I, Baranowska-Bosiacka I, Piotrowska K, Metryka E, Safranow K, Chlubek D. 2018. Influence of acetylcholinesterase inhibitors used in Alzheimer's Disease treatment on the activity of antioxidant enzymes and the concentration of glutathione in THP-1 macrophages under fluoride-induced oxidative stress. *Int J Environ Res Pub Health* 16(1).

Guan ZZ, Shan KR, Xiu J, Long YG. 2005. [Fluorosis on expression of nicotinic acetylcholine receptors in protein and gene levels in human SH-SY5Y neuroblastoma cells]. *Chin J Prev Med* 39: 26-29.

Guan ZZ, Shan KR, Wang YN, Dallner G. 2006. [Changes of lipids and nicotinic receptors in rat brains and pheochromocytoma with fluorosis]. *Chin J Endemiol* 25: 121-124.

Haojun Z, Yaoling W, Ke Z, Jin L, Junling W. 2012. Effects of NaF on the expression of intracellular Ca2+ fluxes and apoptosis and the antagonism of taurine in murine neuron. *Toxicol Mech Methods* 22: 305-308.

Hong-Liang L, Qiang Z, Yu-Shan C, Lei Z, Gang F, Chang-Chun H, Liang Z, Aiguo W. 2014. Fluoride-induced thyroid cell apoptosis. *Fluoride* 47: 161-169.

Inkielewicz-Stepniak I, Radomski MW, Wozniak M. 2012. Fisetin prevents fluoride- and dexamethasone-induced oxidative damage in osteoblast and hippocampal cells. *Food Chem Toxicol* 50: 583-589.

Jin TX, Guan ZZ, Zhang H. 2011. [The effect of fluoride on α subunit of calcium/calmodulin-dependent protein kinase-II mRNA and protein expression in central nervous system]. *Chin J Endemiol* 30: 247-250.

Kariya T, Kotani M, Field JB. 1974. Effects of sodium fluoride and other metabolic inhibitors on basal and TSH stimulated cyclic AMP and thyroid metabolism. *Metab Clin Exper* 23: 967-973.

Ke L, Zheng X, Sun Y, Ouyang W, Zhang Z. 2016. Effects of sodium fluoride on lipid peroxidation and PARP, XBP-1 expression in PC12 cell. *Biol Trace Elem Res* 173: 161-167.

Lee J, Han YE, Favorov O, Tommerdahl M, Whitsel B, Lee CJ. 2016. Fluoride induces a volume reduction in CA1 hippocampal slices via MAP kinase pathway through volume regulated anion channels. *Exp Neurobiol* 25: 72-78.

Levesque L, Mizzen CA, McLachlan DR, Fraser PE. 2000. Ligand specific effects on aluminum incorporation and toxicity in neurons and astrocytes. *Brain Res* 877: 191-202.

Li H, Gao MT, Xu KY, Wang CY. 2007. Effect of sodium fluoride on the primary porcine thyroid cells and thyroid peroxidase activity. *J Clin Rehabil Tissue Eng Res* 11: 7425-7428.

Li H, Gao MT, Xu KY, Cui MY, Dai X. 2008. [Effect of fluoride on thyroid functioning in primary porcine thyrocyte]. *Chin J Endemiol* 27: 38-40.

Li H, Huang H, Xu Y, Gao Y, Liu Z. 2010. [Toxic effects of fluoride on rat cerebral cortex astrocytes in vitro]. *J Hyg Res* 39: 86-88.

Liu H, Zeng Q, Cui Y, Yu L, Zhao L, Hou C, Zhang S, Zhang L, Fu G, Liu Y, Jiang C, Chen X, Wang A. 2014. The effects and underlying mechanism of excessive iodide on excessive fluoride-induced thyroid cytotoxicity. *Environ Toxicol Pharmacol* 38: 332-340.

Liu HL, Zeng Q, Cui YS, Zhao L, Zhang L, Fu G, Hou CC, Zhang S, Yu LY, Jiang CY, Wang ZL, Chen XM, Wang AG. 2014. The role of the IRE1 pathway in excessive iodide- and/or fluoride-induced apoptosis in Nthy-ori 3-1 cells in vitro. *Toxicol Lett* 224: 341-348.

Liu YJ, Guan ZZ, Gao Q, Pei JJ. 2011. Increased level of apoptosis in rat brains and SH-SY5Y cells exposed to excessive fluoride-A mechanism connected with activating JNK phosphorylation. *Toxicol Lett* 204: 183-189.

Liu Y, Gao Q, Tang Z, Zhang X, Guan Z. 2015. [The expression and correlation between neural nicotinic acetylcholine receptor subunit α3 and mitogen-activated protein kinase cell signaling transduction pathway in human neuroblastoma cell line SH-SY5Y overexposed to fluoride]. *Chin J Endemiol* 34: 553-558.

Madaoui S, Rappaport L, Nunez J. 1974. Prostaglandins and in vitro TSH-dependent iodide binding by rat thyroid glands. *Biochimie* 56: 109-113.

Nakagawa-Yagi Y, Saito Y, Kitoh N, Ogane N, Fujisawa E, Nakamura H. 1993. Fluoride causes suppression of neurite outgrowth in human neuroblastoma via an influx of extracellular calcium. *Biochem Biophys Res Commun* 191: 727-736.

Ong J, Kerr DIB. 1995. Interactions of N-ethylmaleimide and aluminium fluoride with GABA(B) receptor function in rat neocortical slices. *Eur J Pharmacol* 287: 197-200.

Pastan I, Macchia V, Katzen R. 1968. Effect of fluoride on the metabolic activity of thyroid slices. *Endocrinology* 83: 157-160.

Rubakhova VM. 1977. [Effect of serotonin and sodium fluoride on visceral nerve conductors]. *Vyestsi Akademii Navuk BSSR Syeryya Biyalahichnykh Navuk* 1: 117-119.

Shayiq RM, Raza H, Kidwai AM. 1986. Fluoride and lipid peroxidation a comparative study in different rat tissues. *Bull Environ Contam Toxicol* 37: 70-76.

Shuhua X, Ziyou L, Ling Y, Fei W, Sun G. 2012. A role of fluoride on free radical generation and oxidative stress in BV-2 microglia cells. *Mediators Inflamm* 2012: 1-8.

Singh P, Das TK. 2019. Ultrastructural localization of 4-hydroxynonenal adducts in fluoride-exposed cells: Protective role of dietary antioxidants. *Fluoride* 52(1): 49-58.

Taylor P. 1972. Comparison of the effects of various agents on thyroidal adenyl cyclase activity with their effects on thyroid hormone release. *J Endocrinol* 54: 137-145.

Tu W, Zhang Q, Liu Y, Han LY, Wang Q, Chen PP, Zhang S, Wang AG, Zhou X. 2018. Fluoride induces apoptosis via inhibiting SIRT1 activity to activate mitochondrial p53 pathway in human neuroblastoma SH-SY5Y cells. *Toxicol Appl Pharmacol* 347: 60-69.

van der Voet GB, Schijns O, de Wolff FA. 1999. Fluoride enhances the effect of aluminium chloride on interconnections between aggregates of hippocampal neurons. *Arch Physiol Biochem* 107: 15-21.

Wang JL. 2007. [Effect of fluoride on the intracellular Ca2+ in neurons of mice]. *Chin J Endemiol* 26: 505-507.

Wang J, Gao Y, Cheng X, Yang J, Zhao Y, Xu H, Zhu Y, Yan Z, Manthari RK, Mehdi OM, Wang J. 2019. GSTO1 acts as a mediator in sodium fluoride-induced alterations of learning and memory related factors expressions in the hippocampus cell line. *Chemosphere* 226: 201-209.

Wei N, Dong YT, Deng J, Wang Y, Qi XL, Yu WF, Xiao Y, Zhou JJ, Guan ZZ. 2018. Changed expressions of N-methyl-d-aspartate receptors in the brains of rats and primary neurons exposed to high level of fluoride. *J Trace Elem Med Biol* 45: 31-40.

Willems CB-V, Sande J, Dumont JE. 1972. Inhibition of thyroid secretion by sodium fluoride in vitro. *Biochim Biophys Acta* 264: 197-204.

Woodward JJ, Harms J. 1992. Potentiation of N-methyl-D-aspartate-stimulated dopamine release from rat brain slices by aluminum fluoride and carbachol. *J Neurochem* 58: 1547-1554.

Wu J, Cheng M, Liu Q, Yang J, Wu S, Lu X, Jin C, Ma H, Cai Y. 2015. Protective role of tert-butylhydroquinone against sodium fluoride-induced oxidative stress and apoptosis in PC12 cells. *Cell Mol Neurobiol* 35: 1017-1025.

Xia T, Zhang M, He WH, He P, Wang AG. 2007. [Effects of fluoride on neural cell adhesion molecules mRNA and protein expression levels in primary rat hippocampal neurons]. *Chin J Prev Med* 41: 475-478.

Xu B, Xu Z, Xia T, He P, Gao P, He W, Zhang M, Guo L, Niu Q, Wang A. 2011. Effects of the Fas/Fas-L pathway on fluoride-induced apoptosis in SH-SY5Y cells. *Environ Toxicol* 26: 86-92.

Xu Z, Xu B, Xia T, He W, Gao P, Guo L, Wang Z, Niu Q, Wang A. 2013. Relationship between intracellular Ca2+ and ROS during fluoride-induced injury in SH-SY5Y cells. *Environ Toxicol* 28: 307-312.

Yamashita K, Field JB. 1972. Elevation of cyclic guanosine 3,5; monophosphate levels in dog thyroid slices caused by acetylcholine and sodium fluoride. *J Biol Chem* 247: 7062-7066.

Yan L, Liu S, Wang C, Wang F, Song Y, Yan N, Xi S, Liu Z, Sun G. 2013. JNK and NADPH oxidase involved in fluoride-induced oxidative stress in BV-2 microglia cells. *Mediators Inflamm* 2013: 895-975.

Zhang CY, Chen R, Wang F, Ren C, Zhang P, Li Q, Li HH, Guo KT, Geng DQ, Liu CF. 2016. EGb-761 attenuates the anti-proliferative activity of fluoride via DDK1 in PC-12 cells. *Neurochem Res* 42(2): 606-614.

Zhang M, Wang A, He W, He P, Xu B, Xia T, Chen X, Yang K. 2007. Effects of fluoride on the expression of NCAM, oxidative stress, and apoptosis in primary cultured hippocampal neurons. *Toxicology* 236: 208-216.

Zhang M, Wang A, Xia T, He P. 2008. Effects of fluoride on DNA damage, S-phase cell-cycle arrest and the expression of NF-kappaB in primary cultured rat hippocampal neurons. *Toxicol Lett* 179: 1-5.

Zhang S, Zheng X, Sun Y, Wang Y, Zhang Z. 2015. Alterations in oxidative stress and apoptosis in cultured PC12 cells exposed to fluoride. *Fluoride* 48: 213-222.

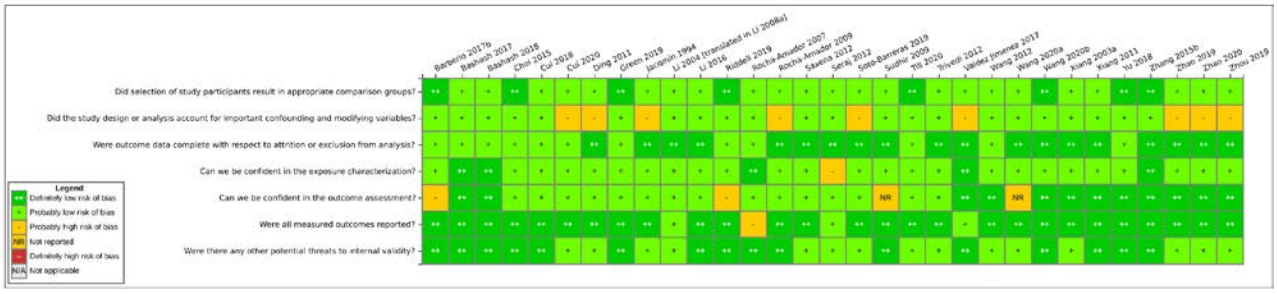
Zhao L, Xiao Y, Deng CM, Tan LC, Guan ZZ. 2016. Protective effect of lovastatin on neurotoxicity of excessive fluoride in primary hippocampal neurons. *Fluoride* 49: 36-46.

Zhao XL, Gao WH, Zhao ZL. 1994. [Effects of sodium fluoride on the activity of Ca2+Mg(2+)-ATPase in synaptic membrane in rat brain]. *Chin J Prev Med* 28: 264-266.

Appendix 3. Risk-of-bias Figures

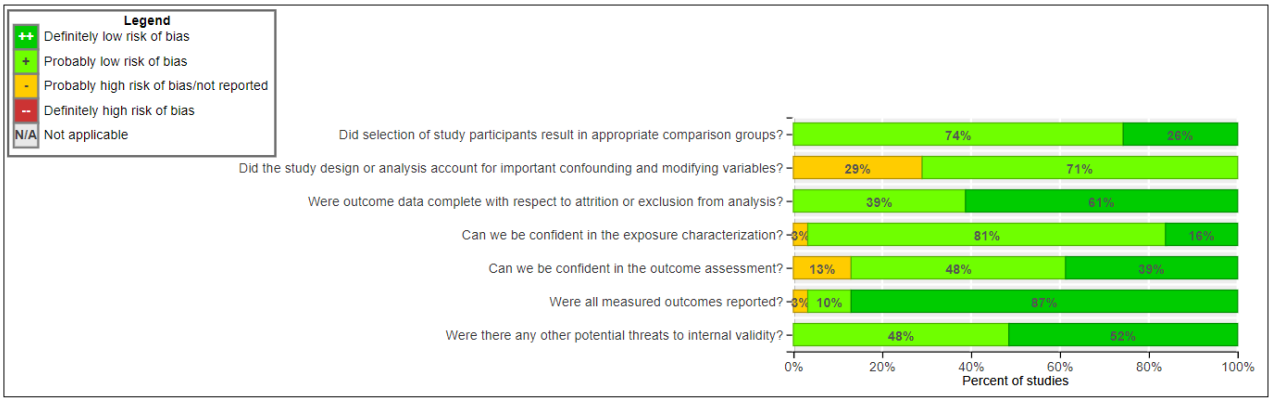
Studies in Humans

Figure A3-1. Risk-of-bias Heatmap for Lower Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure



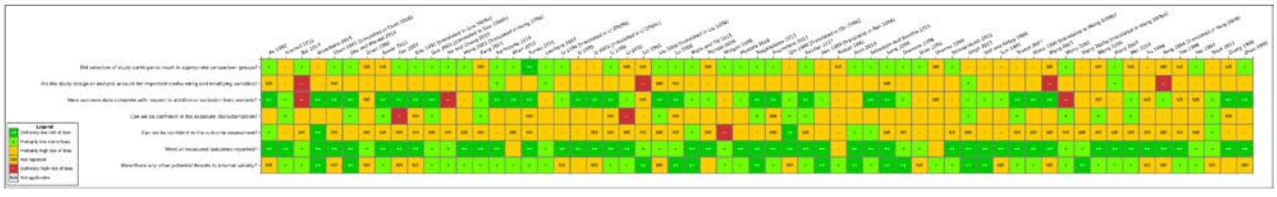
Interactive figure and additional study details in HAWC here.

Figure A3-2. Risk-of-bias Bar Chart for Lower Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure



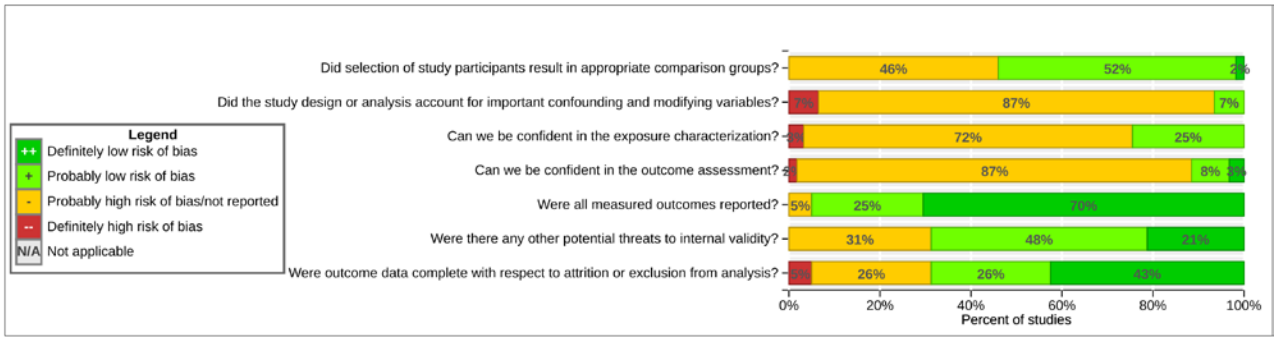
Interactive figure and additional study details in HAWC here.

Figure A3-3. Risk-of-bias Heatmap for Higher Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure



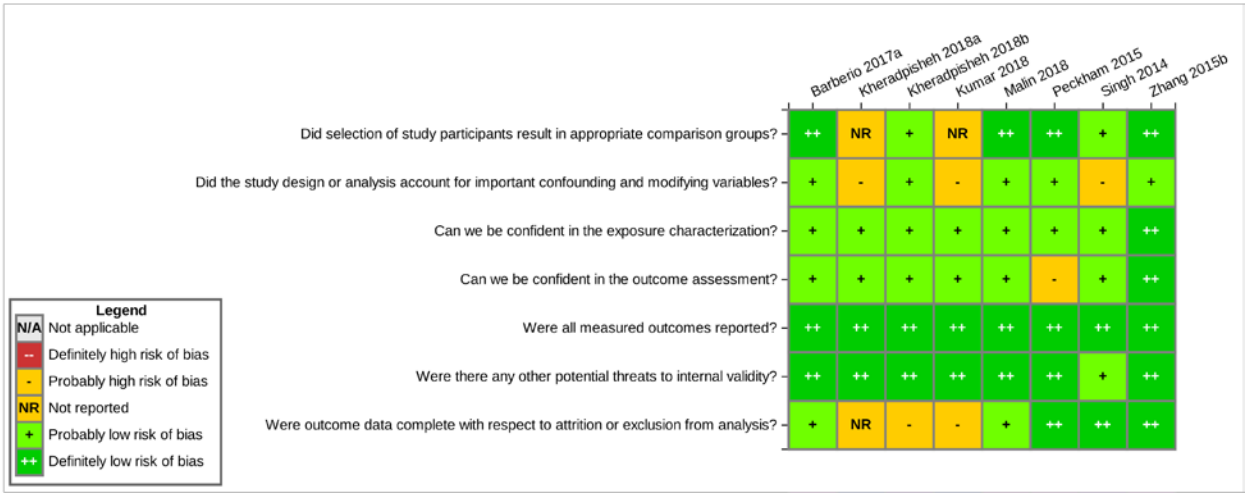
Interactive figure and additional study details in HAWC here.

Figure A3-4. Risk-of-bias Bar Chart for Higher Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure



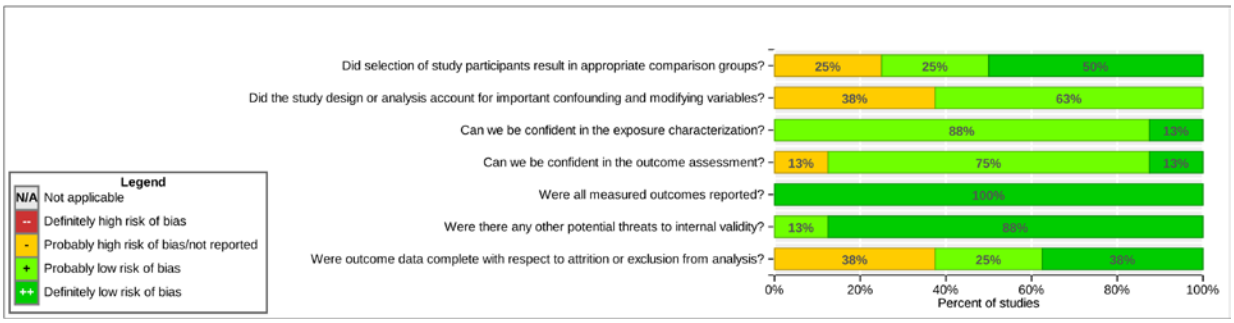
Interactive figure and additional study details in HAWC here.

Figure A3-5. Risk-of-bias Heatmap for Lower Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure



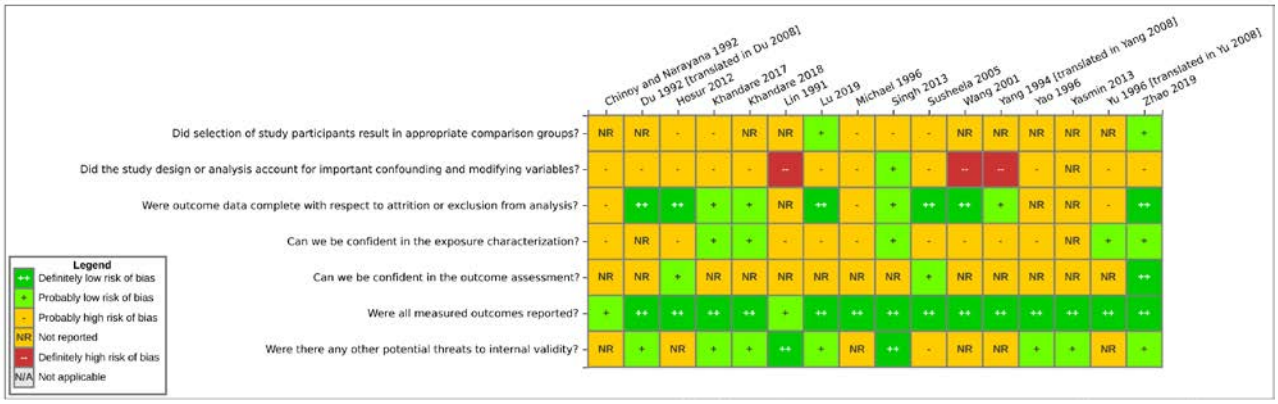
Interactive figure and additional study details in HAWC here.

Figure A3-6. Risk-of-bias Bar Chart for Lower Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure



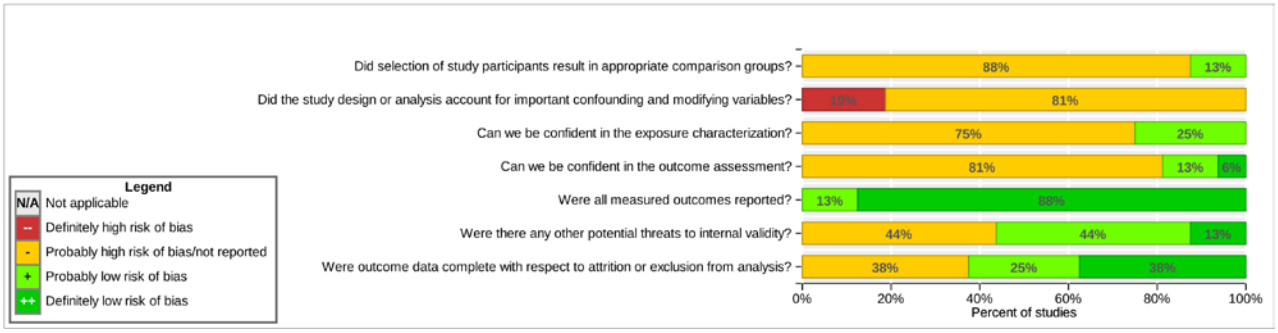
Interactive figure and additional study details in HAWC here.

Figure A3-7. Risk-of-bias Heatmap for Higher Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

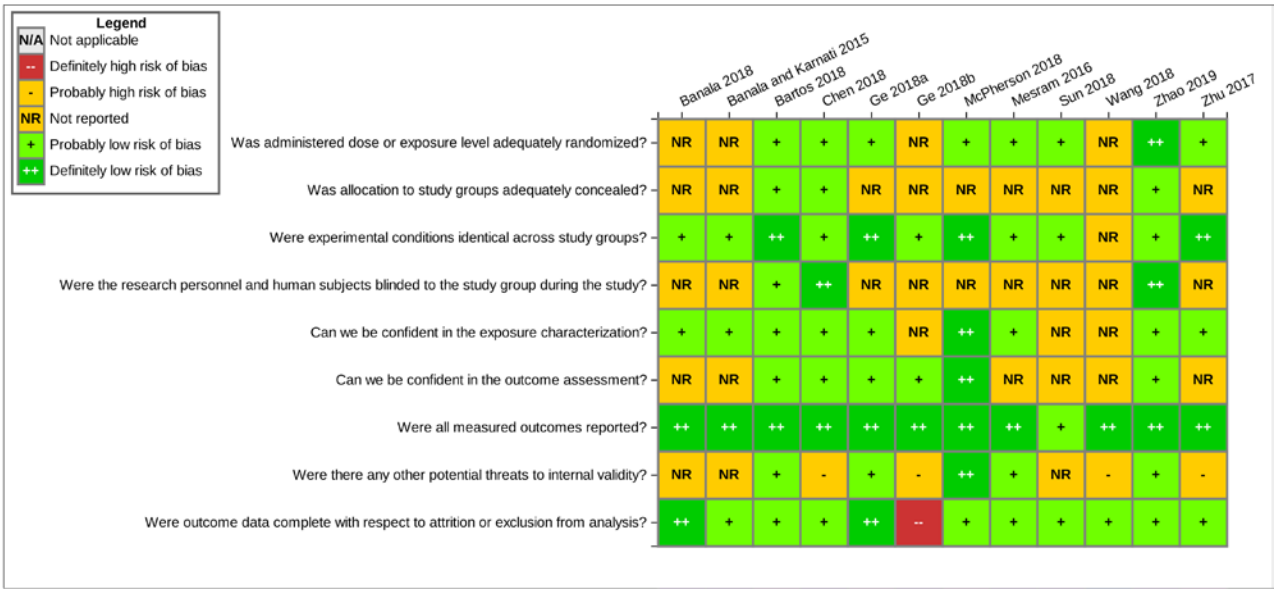
Figure A3-8. Risk-of-bias Bar Chart for Higher Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

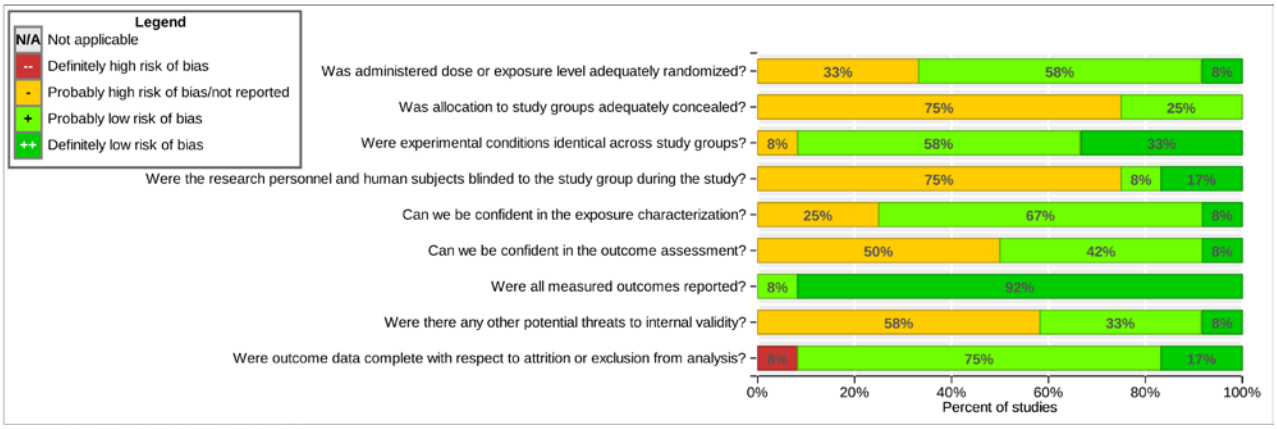
Studies in Non-human Animals

Figure A3-9. Risk-of-bias Heatmap for New Developmental Animal Learning and Memory Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

Figure A3-10. Risk-of-bias Bar Chart for New Developmental Animal Learning and Memory Studies Following Fluoride Exposure

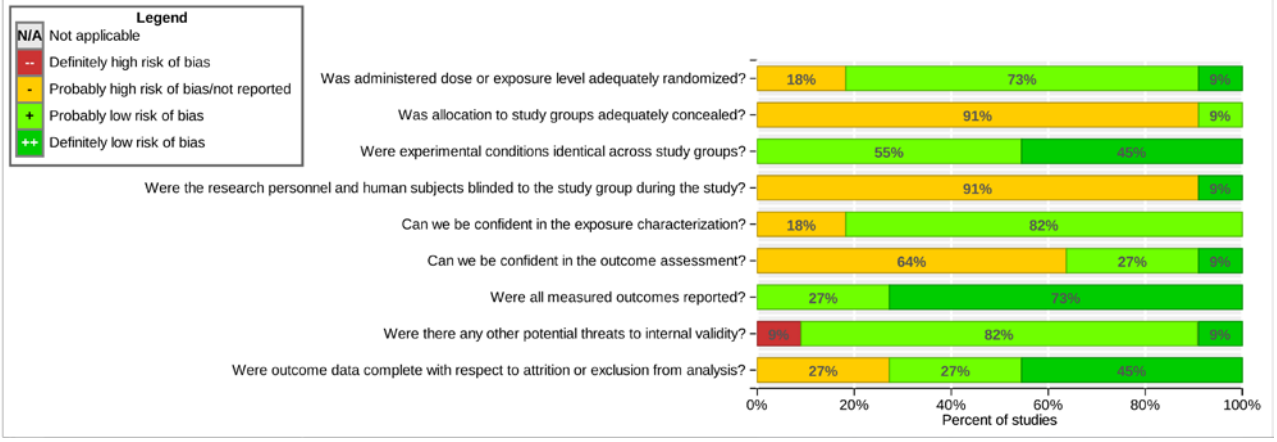


Interactive figure and additional study details in HAWC here.

Figure A3-11. Risk-of-bias Heatmap for New Adult Animal Learning and Memory Studies Following Fluoride Exposure

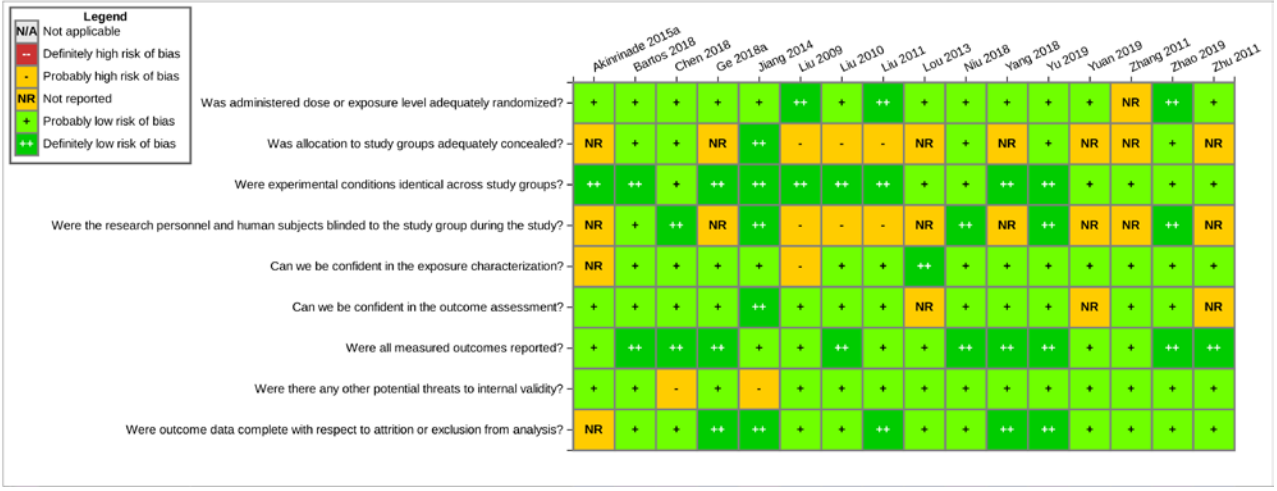
Interactive figure and additional study details in HAWC here.

Figure A3-12. Risk-of-bias Bar Chart for New Adult Animal Learning and Memory Studies Following Fluoride Exposure



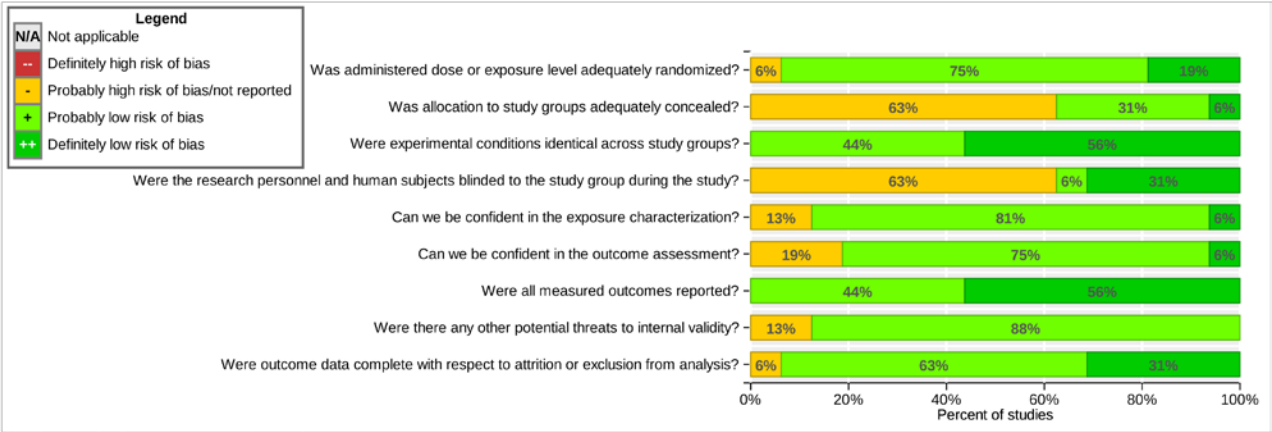
Interactive figure and additional study details in HAWC here.

Figure A3-13. Risk-of-bias Heatmap for Lower Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure



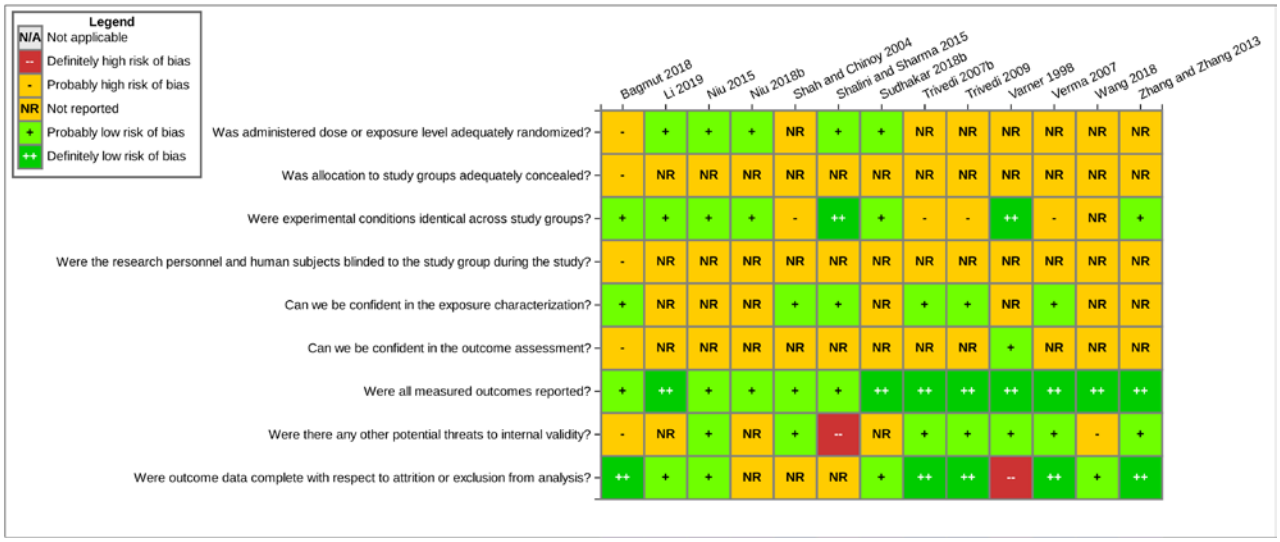
Interactive figure and additional study details in HAWC here.

Figure A3-14. Risk-of-bias Bar Chart for Lower Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure



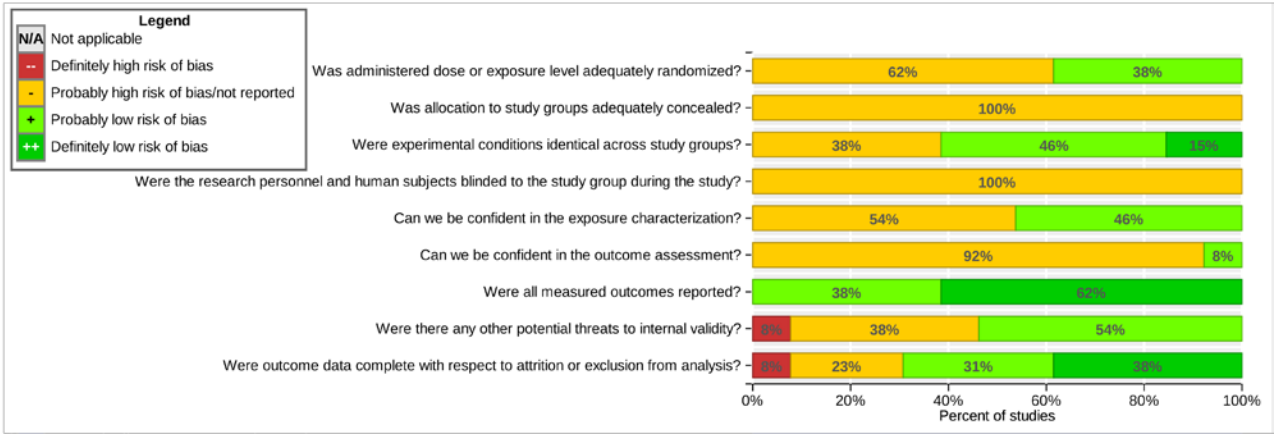
Interactive figure and additional study details in HAWC here.

Figure A3-15. Risk-of-bias Heatmap for Higher Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure



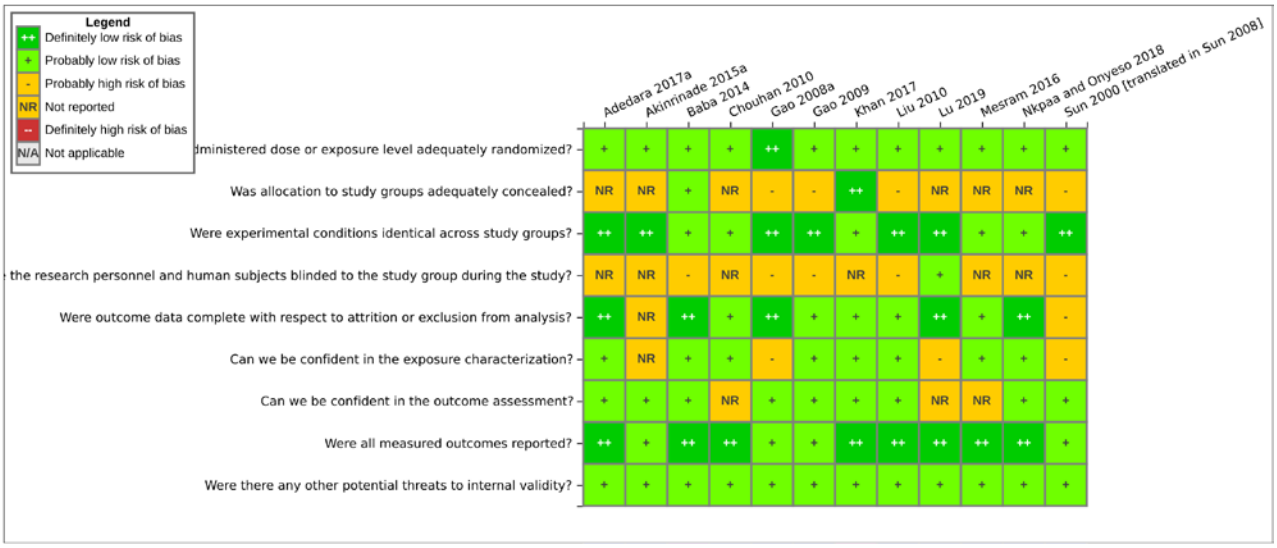
Interactive figure and additional study details in HAWC here.

Figure A3-16. Risk-of-bias Bar Chart for Higher Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure



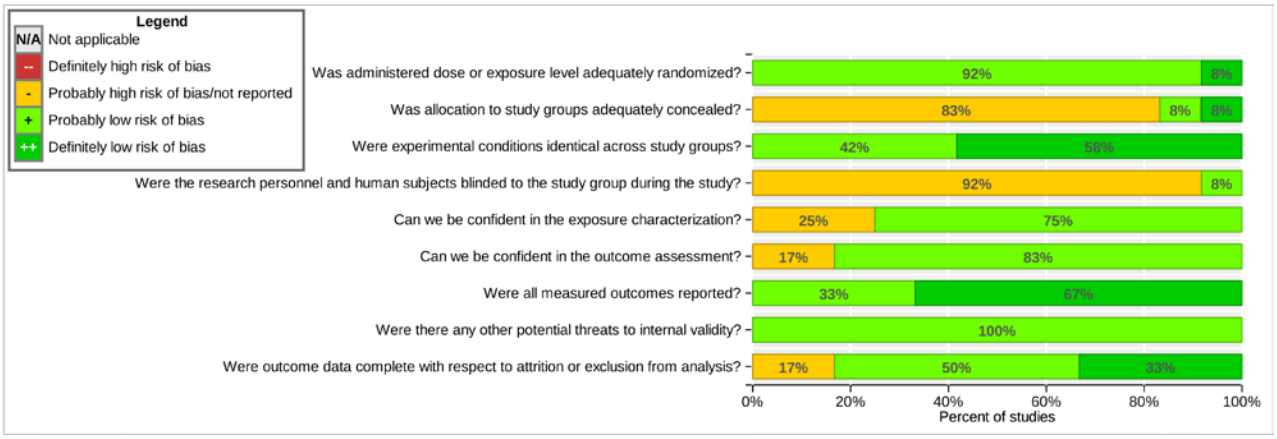
Interactive figure and additional study details in HAWC here.

Figure A3-17. Risk-of-bias Heatmap for Lower Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

Figure A3-18. Risk-of-bias Bar Chart for Lower Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

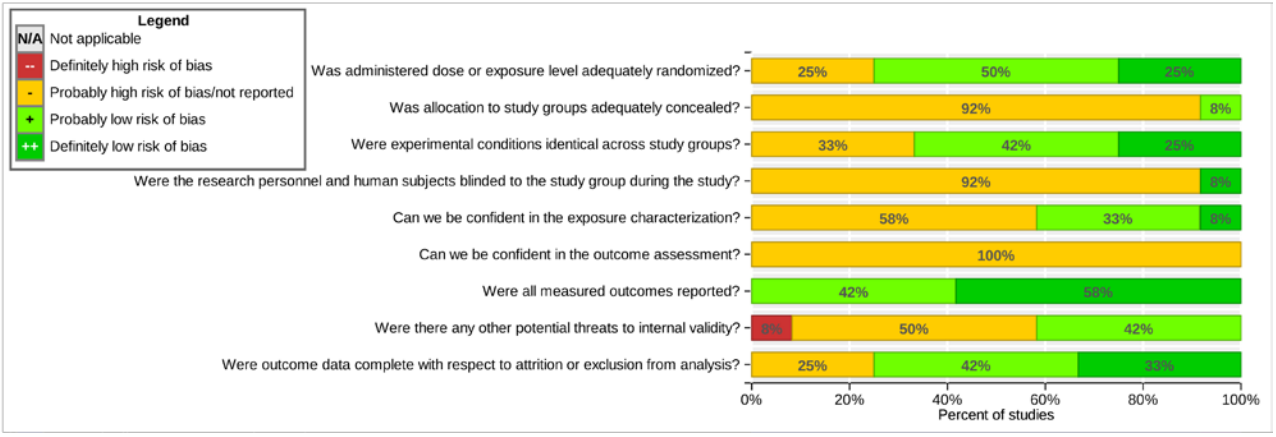


Interactive figure and additional study details in HAWC here.

Figure A3-19. Risk-of-bias Heatmap for Higher Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

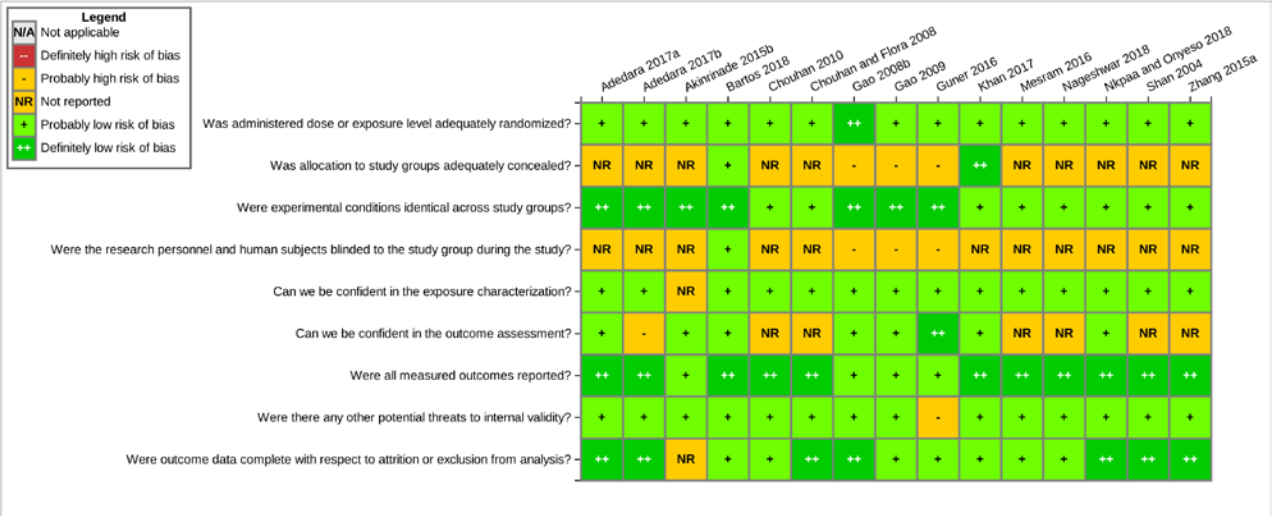
Interactive figure and additional study details in HAWC here.

Figure A3-20. Risk-of-bias Bar Chart for Higher Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure



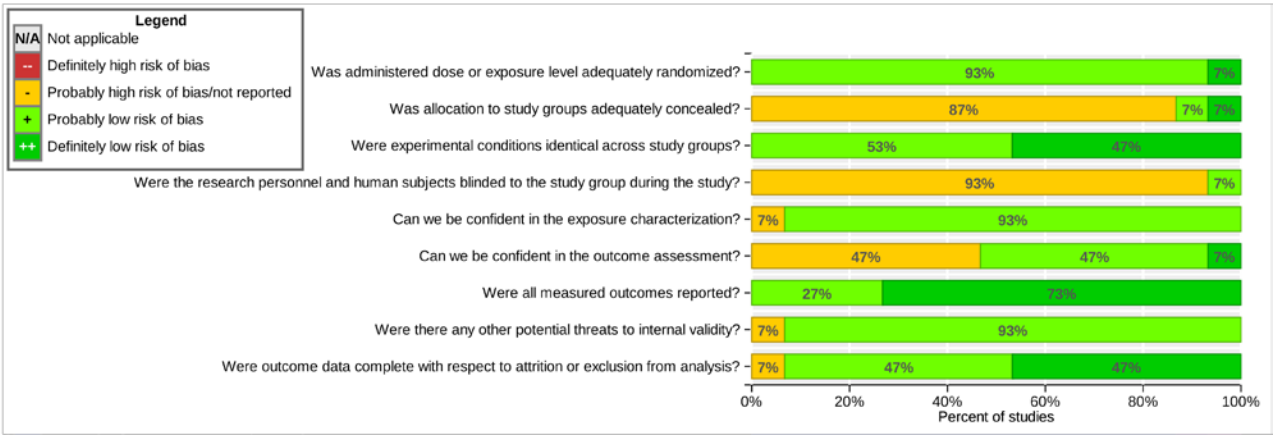
Interactive figure and additional study details in HAWC here.

Figure A3-21. Risk-of-bias Heatmap for Lower Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure



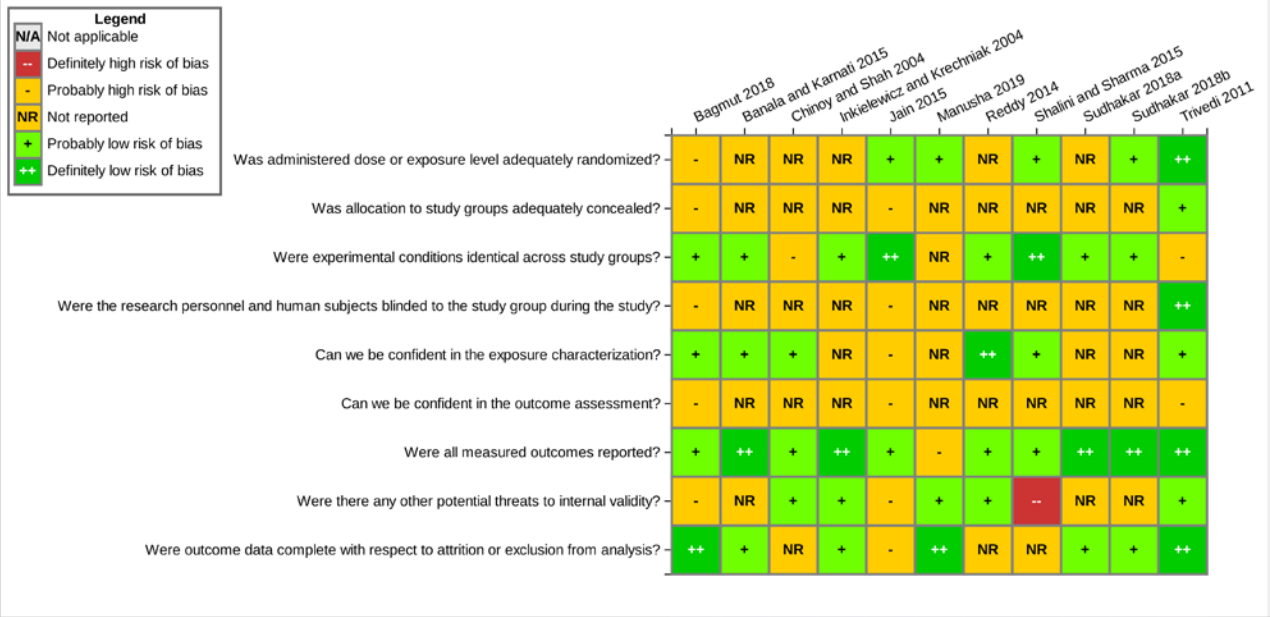
Interactive figure and additional study details in HAWC here.

Figure A3-22. Risk-of-bias Bar Chart for Lower Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure



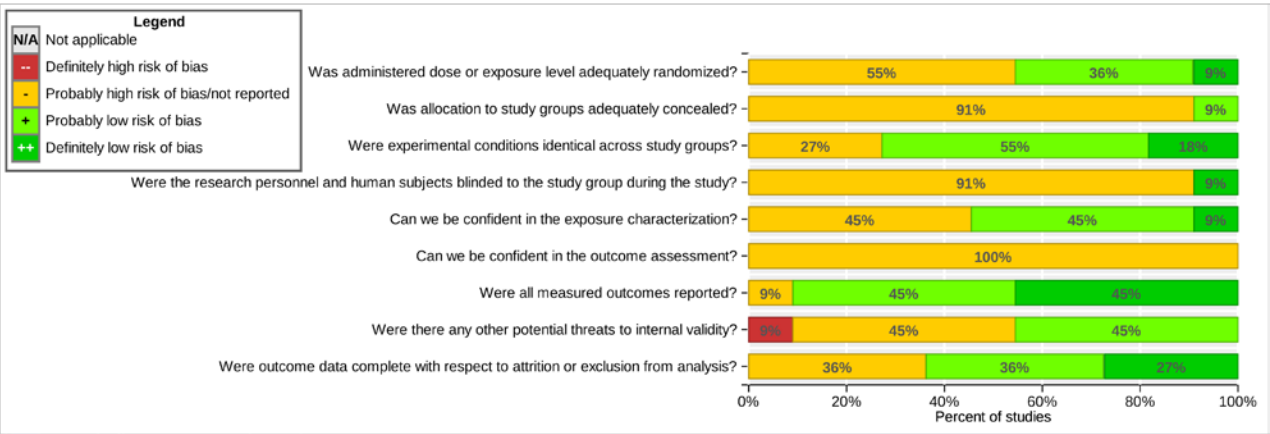
Interactive figure and additional study details in HAWC here.

Figure A3-23. Risk-of-bias Heatmap for Higher Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure



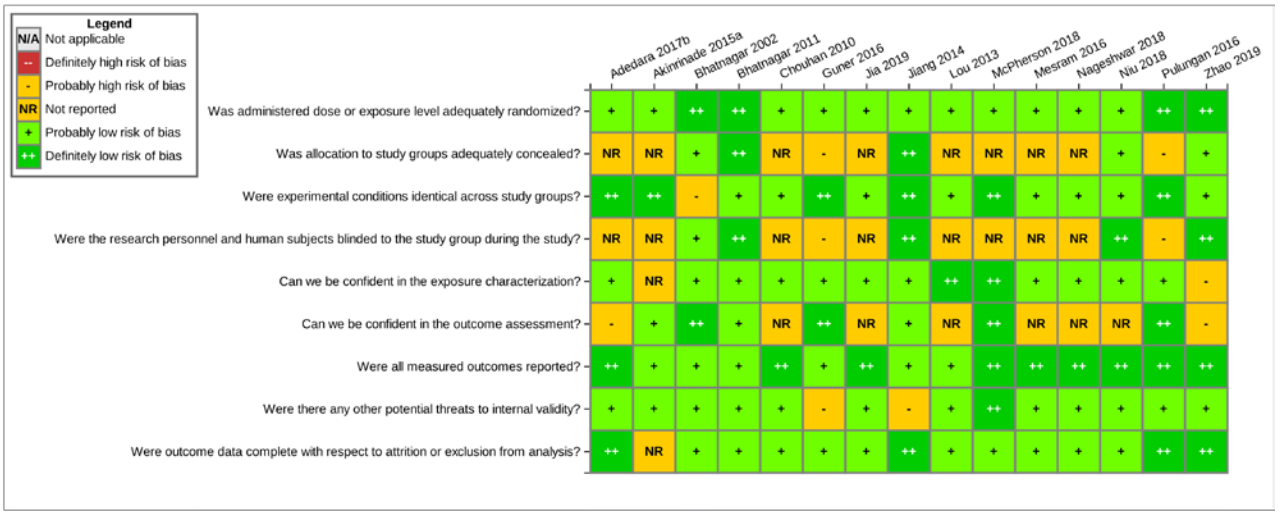
Interactive figure and additional study details in HAWC here.

Figure A3-24. Risk-of-bias Bar Chart for Higher Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure



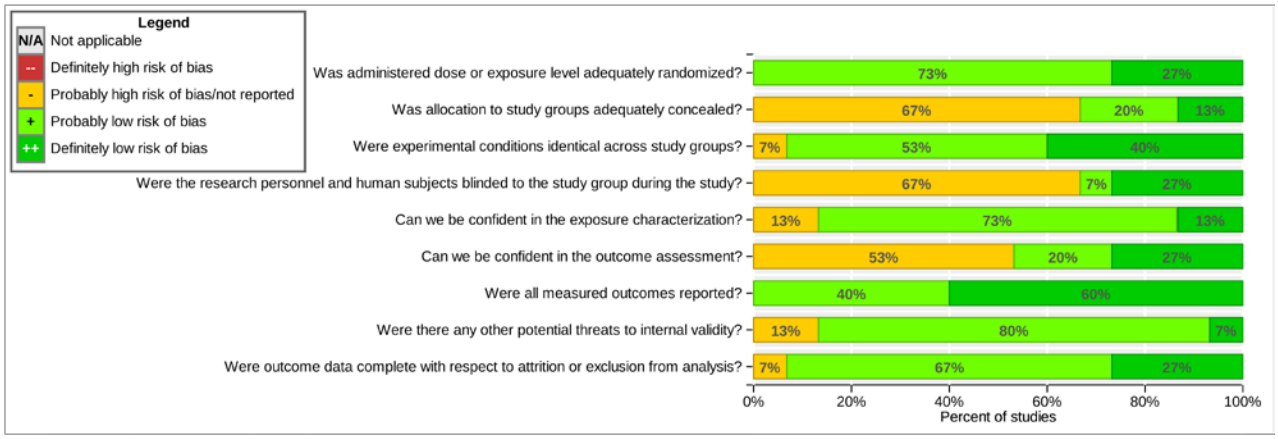
Interactive figure and additional study details in HAWC here.

Figure A3-25. Risk-of-bias Heatmap for Lower Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure



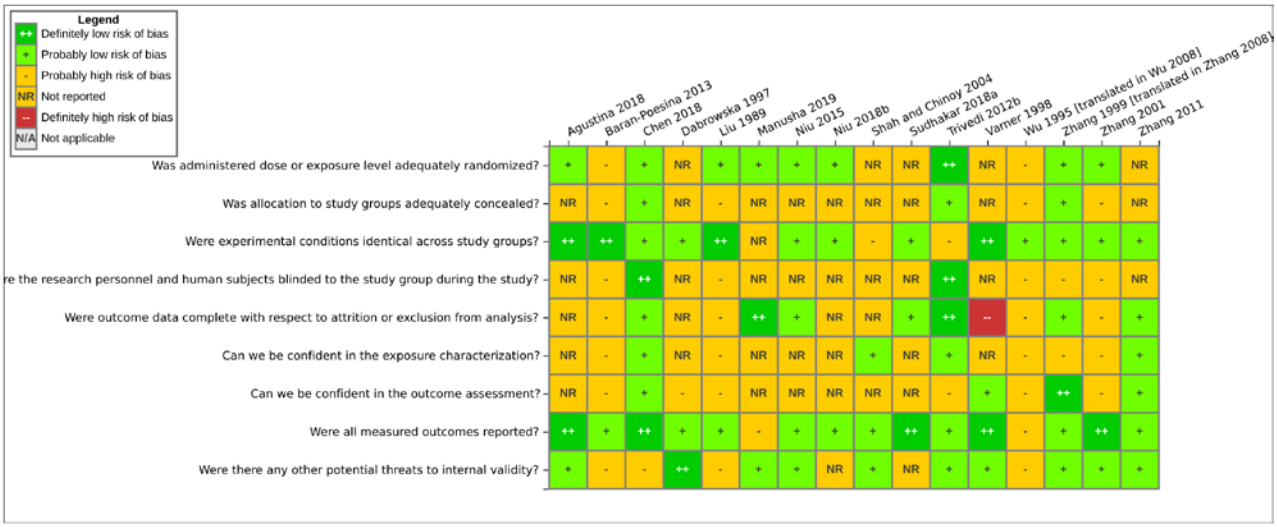
Interactive figure and additional study details in HAWC here.

Figure A3-26. Risk-of-bias Bar Chart for Lower Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure



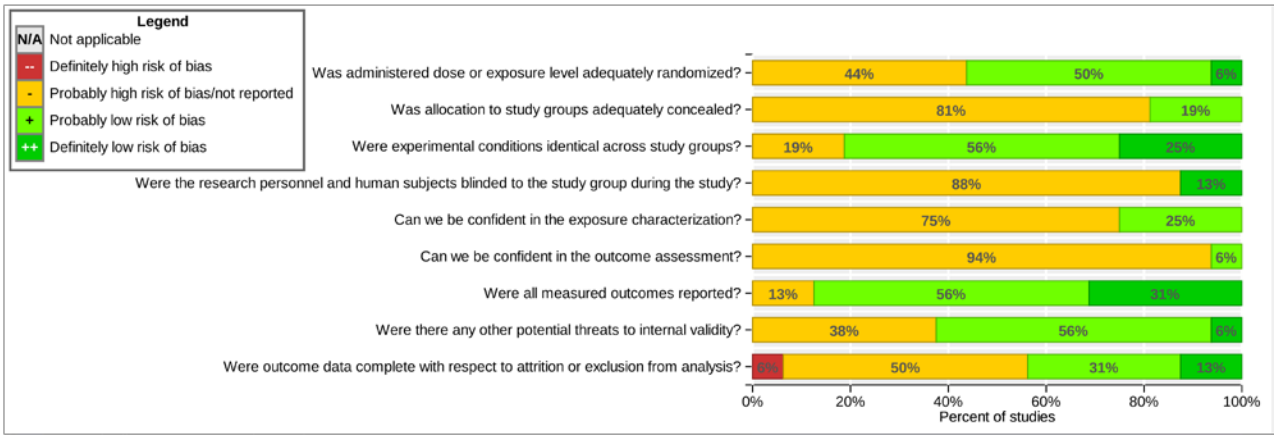
Interactive figure and additional study details in HAWC here.

Figure A3-27. Risk-of-bias Heatmap for Higher Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

Figure A3-28. Risk-of-bias Bar Chart for Higher Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure



Interactive figure and additional study details in HAWC here.

Appendix 4. Details for Lower Risk-of-bias Studies

Barberio *et al.* (2017b)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Canadian Health Measures Survey (cycles 2 and 3) participants (children aged 3−12 years)
- ***Study area:*** general population of Canada
- ***Sample size:*** 2,221 children (1,120 from Cycle 2, 1,101 from Cycle 3)
- ***Data relevant to the review:*** Associations between learning disability or ADHD (Cycle 2 only) assessed by parent or child self-report and urinary fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant increase in adjusted OR for learning disability with unadjusted urinary fluoride (1.02; 95% CI: 1.00, 1.03) when Cycle 2 and 3 were combined. No significant associations with urinary fluoride when adjusted for creatinine and/or specific gravity. No significant association between urinary fluoride and ADHD.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* The comparison groups were selected from Cycles 2 and 3 of the Canadian Health Measures Survey. This is a nationally representative sample of residents living in 10 provinces, with clear exclusion criteria provided. Exclusion only represented about 4% of the target population (all Canadian residents 3−79 years old living in 10 provinces). A table of characteristics of the study population is provided.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the subjects were recruited from the same population using the same methods during the same time frame and exposure groups were similar.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study adjusted for sex, age (3–12 years old), household education, and household income adequacy. Variables to discern fluoride source, including drinking water and dental products, were also considered. Cycle 2 data also included adjustments for: 1) children for whom tap water (vs. bottled or other) was the primary source of drinking water at home or away from home and 2) children who had lived in his or her current home for 3 or more years. Confounders such as parental behavioral and mental health disorders, smoking, and nutrition were not discussed. Co-exposure to lead and arsenic are less likely an issue in this population and the lack of information is not considered to appreciably bias the results.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect*: Not applicable.

  - *Basis for rating:* Probably low risk of bias based on direct evidence that key confounders were addressed and indirect evidence that the methods used to collect the information were valid and reliable and that co-exposures were not an issue.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Covariate data were missing for less than 5% of all analyses, apart from household income; household income was reported for only 71–77% of participants and was imputed for the remainder.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Estimates of urinary fluoride (µmol/L) from spot urine were available for a subsample of respondents. Analysis was performed under standardized operating procedures at the Human Toxicology Laboratory of the Institut National de Santé Publique du Québec (accredited under ISO 17025). Fluoride content of urine samples was analyzed using an Orion pH meter with a fluoride ion-selective electrode with limits of detection of 20 µg/L (Cycle 2) and 10 µg/L (Cycle 3). Urinary dilution was addressed by using creatinine-adjusted levels as well as specific gravity-adjusted levels. In Cycle 3 only, estimates of the fluoride concentration of tap water samples collected from randomly selected households were available. The subsample of households selected for tap water sample collection corresponded to the person-level urine fluoride subsample. Analysis of the fluoride concentration of tap water was performed using a basic anion exchange chromatography procedure, with a limit of detection of 0.006 mg/L. QC methods were not addressed.
    - *Direction/magnitude of effect***:** There is not any specific direction or magnitude of bias expected. Urinary fluoride levels are reflective of a recent exposure. Having a single concurrent measurement may not be reflective of the exposure associated with the outcome, but if subjects lived in the same area throughout life the exposure may be an adequate representation. Although there is possible exposure misclassification it would be non-differential.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* The primary outcome variable, diagnosis of a learning disability by a health professional, was based on a single item from a household survey asked to all respondents: "Do you have a learning disability?". Answer options were: "yes", "no", "don't know", or the participant refused to answer. For Cycle 2, those who indicated having a learning disability were also asked what kind, with the answer options of: "ADD", "ADHD", "dyslexia", or "other". This question was omitted in Cycle 3 and the reason for omission is not described. Parents or guardians answered all questions for children aged 3–11 years, while children 12 years and older answered questions

themselves. The self-reporting of a learning disability did not appear to have been confirmed by medical records or a health professional. (- for methods based on self-report of diagnosis by a health care professional also in Cycle 3 no specific disabilities were described). Blinding was not a concern as spot urine samples were sent to a separate lab and self-reports would not have knowledge of their urine or tap water exposure level (+ for blinding). Overall rating = -.

  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the outcome was measured using an insensitive method in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods sections were reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were appropriate for the study.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include individual exposure measurements and the addressing of potential key confounders but was limited by the cross-sectional study design and insensitive outcome measures.

Bashash *et al.* (2017)

Study Details:

- ***Study design*:** Prospective cohort
- ***Population*:** Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) participants (pregnant mothers and their children aged 4 or 6–12 years).
- ***Study area:*** Mexico City, Mexico
- ***Sample size:*** 299 mother–child pairs, of whom 287 and 211 had data for the general cognitive index (GCI) and IQ analyses, respectively.
- ***Data relevant to the review:*** Adjusted and unadjusted associations between GCI or IQ scores and maternal or child's urinary fluoride concentrations.
- ***Reported association with fluoride exposure:*** Yes: Significant effect between maternal urinary fluoride and IQ score (adjusted β = −2.50; 95% CI: −4.12, −0.59) and GCI score (adjusted β = −3.15; 95% CI: −5.42, −0.87). No significant effects associated with children's urine.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Study participants were selected from two different cohorts from three hospitals in Mexico City that serve low-to-moderate income populations. One cohort was from an observational study of prenatal lead exposure and neurodevelopment outcomes and the other was from a randomized trial of the effect of calcium on maternal blood lead levels. The authors state that participants had no history of psychiatric disorders, high-risk pregnancies, gestational diabetes, illegal drug use, or continuous prescription drugs, but they do not include any information on smoking habits. Study populations appear to be similar, but there may be some differences because subjects were selected from two different cohorts that were recruited from slightly different time periods.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar despite the subjects coming from different original study populations where different methods were used for recruitment.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Data were collected via questionnaire on maternal age, education, marital status at first prenatal visit, birth order, birth weight, gestational age at delivery, maternal smoking, maternal IQ, and HOME scores. All models were adjusted for gestational age at birth, child's sex, birth weight, birth order, child's age at testing, maternal marital status, smoking history, age at delivery, maternal IQ, education, and cohort, with additional testing for children's urinary fluoride, mercury, lead, and calcium. Sensitivity analyses additionally adjusted for HOME score. Confounders not considered included BMI, iodine deficiency, arsenic, and maternal mental health and nutrition. Arsenic is assumed not to be a potential co-exposure in this population as the study authors did not discuss it as an issue but did discuss other co-exposures. Arsenic may have been included in the water quality control program in Mexico City.
  - *Potentially important study-specific confounders:* All key confounders were addressed.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that key confounders including other potential co-exposures were addressed and indirect evidence that the methods used to collect the information were valid and reliable and that arsenic is not likely to be an issue in this study population.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Although there was a large amount of attrition, the study authors clearly describe all reasons for attrition and also provide characteristics to compare those participants included to those excluded. There were some slight differences between those included and those excluded, but there is nothing to indicate that the attrition would potentially bias the results.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Urinary fluoride concentrations were determined in spot urine samples (2nd morning void) collected from mothers (during at least one trimester) and children ages 6–12 years. Fluoride content was measured using ion-selective electrode-based assays. QC methods were described including between laboratory correlations. All samples were measured in duplicate. Extreme outliers were excluded. Urinary dilution was addressed by using creatinine-adjusted levels.
    - *Direction/magnitude of effect*: Not applicable.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Outcome was assessed using the McCarthy Scales of Children's Abilities (MSCA) in 4-year-old children (translated into Spanish) and the Wechsler Abbreviated Scale of Intelligence (WASI) in 6–12-year-olds. The WASI is a well-established test and the validity of both tests is well documented by the authors. Inter-examiner reliability was evaluated and reported with a correlation of 0.99 (++ for methods). The study report stated that psychologists were blind to the children's fluoride exposure (++ for blinding). Overall rating for methods and blinding = ++.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, outcome blindly assessed, and the prospective cohort study design.

Bashash *et al.* (2018)

Study Details:

- ***Study design*:** Prospective cohort
- ***Population*:** ELEMENT participants (pregnant mothers and their children aged 6–12 years)
- ***Study area:*** Mexico City, Mexico
- ***Sample size:*** 210 mother–child pairs
- ***Data relevant to the review:*** Associations between ADHD and other attention/impulsivity scores and maternal urinary fluoride concentrations.
- ***Reported association with fluoride exposure:*** Yes: Significant associations between maternal urinary fluoride and Conners' Rating Scales-Revised (CRS-R) scores, including Cognitive Problems + Inattention Index (adjusted β = 2.54; 95% CI: 0.44, 4.63), DSM-IV Inattention Index (adjusted β = 2.84; 95% CI: 0.84, 4.84), DSM-IV ADHD Total Index (adjusted β = 2.38; 95% CI: 0.42, 4.34), and ADHD Index (adjusted β = 2.47; 95% CI: 0.43, 4.50).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Participants were a subset of mother-child dyads enrolled in various longitudinal birth cohort studies of the Early Life Exposure in Mexico to Environmental Toxicants (ELEMENT) project. Subjects were included from two of the four cohorts for which maternal urinary samples were available. Participants in cohort 2A were recruited between 1997 and 1999, and participants in cohort 3 were recruited from 2001 to 2003. Inclusion and exclusion criteria were applied consistently across the two cohorts. A table of subject characteristics was provided in the study and any differences were considered in the analysis. Study populations appear to be similar, but there may be some differences because subjects were selected from two different cohorts: one from an observational study on prenatal lead exposure and the other from a randomized trial on the effects of calcium on blood lead levels. In addition, they were recruited from slightly different time periods.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposed groups were similar, and any differences were taken into account in the analysis.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Questionnaires were used to collect information on maternal age, maternal education, history of smoking, and marital status during the first pregnancy visit. Child information at birth included birth weight, sex, birth order, and gestational age as calculated by the nurse. Mothers also responded to an SES questionnaire during the visit when the psychometric tests were administered. The Home Observation for Measurement of the Environment (HOME) score was evaluated in a subset of participants. Covariates were selected a priori. Models adjusted for maternal age at

delivery, years of education, marital status, smoking history, gestational age at birth, age at outcome assessment, child's sex, birth order, SES, cohort, and calcium intervention.

- *Potentially important study-specific confounders:* None identified, although this study did not specifically address arsenic or other co-exposures. Bashash *et al.* (2017) addressed potential co-exposure to lead and mercury but did not address arsenic. Arsenic was potentially addressed as part of the water quality program in Mexico City.
  - *Direction/magnitude of effect*: Not applicable.
- *Basis for rating:* Probably low risk of bias based on direct evidence that key confounders were addressed and indirect evidence that the methods used to collect the information were valid and reliable and that arsenic and other potential co-exposures are not likely to be an issue in this study population.

- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Although there was a large amount of attrition from the original cohorts, it was unlikely related to outcome or exposure and there were very little missing data from those included in the study. Of the 231 mothers with a minimum of one maternal urine fluoride measurement and matching outcome identified for the project, only 17 were excluded based on incomplete demographic and outcome information.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Mothers provided at least one spot urine sample during pregnancy. As described in Bashash *et al.* (2017), urinary concentrations were determined on second morning void. Fluoride content was measured using ion-selective electrode-based assay. Bashash *et al.* (2017) describes QC methods. All samples were measured in duplicate and extreme outliers were excluded. Urinary dilution was addressed by using creatinine-adjusted levels.
    - *Direction/magnitude of effect*: N/A
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Behaviors associated with ADHD were assessed using the Spanish version of the Conners' Rating Scales-Revised, which has been validated for the evaluation of ADHD. Mothers completed the CRS-R at the same follow-up visit that the child completed the CPT-II tests. All tests were applied under the supervision of an experienced psychologist (++ for methods); however, a limitation of the study noted by the authors was only using parent reports and not teacher reports as they can vary from one another. Blinding was not reported, but it is unlikely that the mothers were aware of their urinary fluoride levels. Although mothers may have had knowledge that they were receiving fluoride through fluoridated salt or naturally occurring fluoride in their water, they would not have knowledge that this was relevant to the study purpose as

the ADHD tests were conducted for the original cohort (as was acknowledged by the study authors in the discussion). (++ for blinding). Overall rating = ++.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported in sufficient detail.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:*
        - *Statistical analyses:* Statistical analyses used were appropriate for the study.
        - *Other potential concerns:* None identified.
    - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, outcome blindly assessed, and the prospective cohort study design.

Choi *et al.* (2015)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** First grade children (ages 6–8 years)
- ***Study area:*** Mianning County in southern Sichuan, China
- ***Sample size:*** 51 first grade children
- ***Data relevant to the review:*** Associations between learning, memory, IQ (digit span for auditory span and working memory and block design for visual organization and reasoning components of WISC-IV only), visual motor ability, motor ability, and manual dexterity with continuous urine or drinking water fluoride levels. Study also had information based on dental fluorosis score.
- ***Reported association with fluoride exposure:*** Yes: Compared to the normal/questionable dental fluorosis, the moderate/severe dental fluorosis group was associated with significantly lower total (adjusted β = −4.28; 95% CI: −8.22, −0.33) and backward digit span scores (adjusted β = −2.13; 95% CI: −4.24, −0.02). Linear correlation between fluoride in urine (adjusted β = −1.67; 95% CI: −5.46, 2.12) and in drinking water (adjusted β = −1.39; 95% CI: −6.76, 3.98) with total digit span was observed but not significant. Other outcomes not significantly associated with fluoride exposure.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Subjects were selected during the same time frame using the same methods. Fifty-one first-grade children residing in Mianning County in southern Sichuan, China were included in this pilot study. It is not specified if the 51 children represented all the first-grade children from this area or if some refused to participate. Children who did not speak Chinese, were not students at the Primary School of Sunshui Village in Mianning County, or those with chronic or acute disease that might affect neurobehavioral function tests were excluded. Demographic characteristics are presented in Table 1 of the study, which indicates that subjects were similar. Potential confounders are adjusted for in the statistical analyses.
  - *Basis for Rating:* Definitely low risk of bias based on direct evidence that the exposure groups were similar and were recruited within the same time frame using the same methods with no evidence of differences in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The parents or guardians completed a questionnaire on demographic and personal characteristics of the children (sex, age at testing, parity, illnesses before age 3, and past medical history) and caretakers (age, parity, education and occupational histories, residential history, and household income). A 20-µL capillary blood sample was collected at the school by a Mianning County Center for Disease Control (CDC) health practitioner and tested for possible iron deficiency which could be used as a covariate of neurodevelopmental performance. Confounders that were not assessed include: maternal BMI, parental mental health, maternal smoking status, maternal reproductive factors, parental IQ, and HOME score. However, the study authors noted that confounding bias appeared to be limited due to the minimal diversity in the social characteristics of the subjects. The study authors indicated that CDC records documented that levels of other contaminants including arsenic and lead were very low in the area. Iodine differences were not specifically addressed, but there is no indication from the information provided that this might be a concern.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is direct evidence that the key confounders are taken into account and indirect evidence that co-exposure to arsenic is likely not an issue in this area and that methods used for collecting the information were valid and reliable.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The majority of results were reported for the 51 children stated to be included in the pilot study. In Table 5 of the study, the N for each dental fluorosis

category only totals 43, but the text indicates 8 children did not have a Dean Index because teeth had not erupted.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.

- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study used three different measurements of fluoride exposure: well water fluoride concentrations from the residence during pregnancy and onwards, fluoride concentrations from children's first morning urine samples, and degree of children's dental fluorosis. Fluoride concentrations in community well water were measured and recorded by Mianning County CDC; specific methods were not reported, but they likely used standard methods as they were conducted by the CDC and were likely the same as those used to measure the fluoride in urine. Migration of subjects was noted to be limited. Well water fluoride concentrations of the mother's residence during pregnancy and onward were used to characterize a child's lifetime exposure. To provide a measure of the accumulated body burden, each child was given a 330-mL (11.2-oz) bottle of Robust© distilled water (free from fluoride and other contaminants) to drink the night before the clinical examinations, after emptying the bladder and before bedtime. The first urine sample the following morning was collected at home, and the fluoride concentration was determined on a 5-mL sample using an ion-specific electrode at the Mianning CDC. There is no indication that urinary fluoride levels accounted for dilution nor was it clear that the method of administering water to the children and collection methods sufficiently controlled for differences in dilution. One of the investigators, a dentist, performed a blinded dental examination on each child's permanent teeth to rate the degree of dental fluorosis using the Dean Index. The Dean Index is a commonly used index in epidemiological studies and remains the gold standard in the dentistry armamentarium. The Index has the following classifications: normal, questionable, very mild, mild, moderate, and severe. Quality control (QC) procedures are not reported but were likely appropriate.
    - *Direction/magnitude of effect***:** Current levels were used to assess lifetime exposure. This is likely to be a non-differential exposure misclassification and direction of bias is unknown. Because subject migration appears to be limited, it is likely that the current fluoride levels are adequate reflections of past exposure. Dental fluorosis would be an indicator that exposure occurred in the past and there was a fair correlation between degree of dental fluorosis and current urine and water fluoride levels, with both increasing with increasing levels of dental fluorosis.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measure exposure.

- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study authors adopted culture-independent tests considered feasible for children aged 6 to 8 years. The Wide Range Assessment of Memory and Learning

(WRAML) was used for the assessment of memory and learning. Three subtests were also used. The Finger Windows subtest assesses sequential visual memory. The Design Memory subtest assesses the ability to reproduce designs from memory following a brief exposure. The Visual Learning subtest assesses the ability to learn the locations of pictured objects over repeated exposures. The Wechsler Intelligence Scale for Children-Revised (WISC-IV) included digit span for auditory span and working memory and block design for visual organization and reasoning. The grooved pegboard test assesses manual dexterity. The tests used have been validated on a western population. Although there is no information provided to indicate that they were validated on the study population, the study authors indicated that the tests were culture-independent (+ for methods). Blinding of the outcome assessors or steps to minimize potential bias was not reported. However, it is unlikely that the assessors had knowledge of the individual exposure as children all came from the same area, and water and urine levels were tested at the CDC. (+ for blinding). Overall = +.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that all outcomes were assessed blindly using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were appropriate. Data were log-transformed when necessary.
    - *Other potential concerns:* It should be noted that this study was a pilot study and, therefore, had a relatively small sample size (i.e., 51 children).
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in the confounding, exposure, and outcome risk-of-bias domains. Study strengths include individual fluoride measurements with blinding at outcome assessment likely. All key confounders and many other confounders were taken into account in the study design or analysis.

Cui *et al.* (2018)

Study Details:

- ***Study design*:** Cross-sectional

- ***Population*:** School children aged 7–12 years from four schools in two districts in China with different fluoride levels
- ***Study area:*** Jinghai and Dagang in Tianjin City, China
- ***Sample size:*** 323 school children
- ***Data relevant to the review:*** IQ scores by urine fluoride levels.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between IQ score and urinary fluoride (adjusted β = −2.47).

Risk of Bias:

- **Author contacts:**
  - Authors were contacted in June 2019 to obtain additional information for risk-of-bias evaluation.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Four schools were selected from the same district in China. The schools were selected based on levels of fluoride in the local drinking water and the degree of school cooperation. No details were provided on the number of schools in given areas or the difficulty in getting school cooperation. It was noted that the residents in the four areas had similar living habits, economic situations, and educational standards. Although authors do not provide the specific data to support this, fluoride levels and IQ scores were provided by different subject characteristics. The areas were classified as historically endemic fluorosis and non-fluorosis. Cluster sampling was used to select the grades in each school according to previously set child ages, and classroom was randomly selected with all students within a selected classroom included. Reasons for exclusion do not appear to be related to exposure or outcome.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and recruited within the same time frame using the same methods, with no evidence of differences in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The measurements of all covariates were obtained by structured questionnaires that were completed by children with the help of their parents. Confounders that were assessed include: child's gender, child's ethnicity, child's age, child's BMI, birth (normal vs abnormal), mother's age at delivery, mother's education, income per family member, mother's smoking/alcohol during pregnancy, family member smoking, environmental noise, iodine region (non-endemic vs iodine-excess-endemic area), factory within 30 m of residence, iodine salt, diet supplements, seafood/pickled food/tea consumption, surface water consumption, physical activity, stress, anger, anxiety/depression, trauma, having a cold 5 times a year, thyroid disease in relatives, mental retardation in relatives, and cancer in relatives. Covariates not considered include parity, maternal and paternal IQ, and quantity and quality of caregiving environment (e.g., HOME score). The authors report that there are no other environmentally toxic substances that may affect intelligence, such as high arsenic or iodine deficiency according to the Tianjin Centers for Disease Prevention and Control.

  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect*: Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key confounders are considered, methods for collecting the information are valid and reliable, and co-exposure to arsenic is likely not an issue in this area.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of the 400 children enrolled, 35 were excluded because they did not have informed consent signed by a guardian or they moved out of the area. Forty-two children were excluded because they did not have a DRD2 genotyping measurement. It is unclear if these children were from the same schools or if they were evenly distributed throughout the study area. It was also unclear if the excluded subjects were similar to those included in the study. In the study, some analyses had fewer than the 323 subjects, but this seems reasonable given the subgroups that were being evaluated.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Although children were selected based on area fluoride levels, fluoride in the urine was used in the analysis. Urine was collected from each child the morning of enrollment and analyzed within a week. Fluoride levels were measured using an ion-selective electrode according to the China standard. A brief description of the method was provided, but no QC methods were reported. The study authors did not account for urinary dilution in the spot samples.
    - *Direction/magnitude of effect*: Not accounting for dilution could cause there to be some exposure misclassification. The direction and magnitude would depend on where the differences occurred.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using acceptable methods that provide individual levels of exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* IQ was measured by professionals using the Combined Raven's Test-The Rural in China method, which is the appropriate test for the study population (++ for methods). Blinding or other methods to reduce bias were not reported. Although it is unlikely that the outcome assessor would have knowledge of the child's urine fluoride levels, there is potential that they would know if the child was from an endemic or non-endemic area if the IQ tests were conducted at the child's school, and there was no information provided on how the IQ tests were administered. Correspondence with the study author noted the cross-sectional nature of the study with outcome and exposure assessed at the same time making the outcome assessors blind to the exposure. However, there is still potential for knowledge of the area (+ for blinding).

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes in the abstract, introduction, and methods are reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were appropriate. Data were log-transformed when necessary.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements but is limited by the cross-sectional study design and lack of accounting for urine dilution. All key confounders were accounted for in the study design or analysis.

Cui *et al.* (2020)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** School children aged 7–12 years
- ***Study area:*** Tianjin City, China (one randomly selected school from each district based on iodine levels in the water), presumably was an expansion of the Cui *et al.* (2018)
- ***Sample size:*** 498 school children
- ***Data relevant to the review:*** IQ scores by urine fluoride levels.
- ***Reported association with fluoride exposure:*** No: No significant difference in IQ score based on a one-way ANOVA in the three different urinary fluoride categories.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for the 2020 publication. Authors were contacted in June 2019 for additional information on the Cui *et al.* (2018) publication. Information obtained from that correspondence may have been used for additional information in the 2020 publication.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* Subjects were recruited from 2014 to 2018. One school was selected from each district where water concentrations of water iodine were <10, 10–100, 100–150, 150–300 and >300 µg/L. In each school, classes were randomly sampled for the appropriate age group of 7–12 years old. A table of subject characteristics was provided by IQ. A total of 620 children were recruited, and 122 children who did not have complete information or enough blood sample were excluded. Reasons for exclusion do not appear to be related to exposure or outcome. The characteristics of the 498 included children are presented.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited within the same time frame using the same methods, with no evidence of differences in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* It was noted by the study authors that there were no other environmental poisons except water fluoride. Other studies also conducted in this area of China noted specifically that arsenic was not a concern. Iodine was addressed as that was one of the main points of the study. Twenty-one factors (provided in Table 1 of the study) were selected as confounders, and a homemade questionnaire of unspecified validity was used for obtaining the information. It was noted that child age, stress, and anger were significantly associated with IQ although it is unclear if these varied by fluoride level. However, Cui *et al.* (2018) indicates that stress and anger were not significantly associated with fluoride, and it is assumed that results would be similar for this study even though more children were included in the current study.
  - *Potentially important study-specific confounders:* Age (children 7–12 years old)
    - *Direction/magnitude of effect*: Age is a potential confounder for IQ, even in the narrow age range evaluated in this study. The direction of effects may depend on the number of children in each age group within the different urinary fluoride categories; however, these data were not provided. In general, there were fewer subjects ≤ 9 years of age (i.e., 111) compared to > 9 years of age (i.e., 387) with a significantly higher IQ in the ≤9-year-old age group. Therefore, if exposure were higher in the older subjects, this could bias away from the null.
  - *Basis for rating:* Probably high risk of bias because there is indirect evidence that age was not addressed as a confounder and it may be related to both IQ and exposure.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of the 620 (20%) children recruited, 122 were excluded due to incomplete information or inadequate blood sample. No information was provided to indicate if there were similarities or differences in the children included versus the children excluded, but exclusion is unlikely to be related to either outcome or exposure.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* Children's morning urine was collected with a clean polyethylene tube and fluoride was measured using a fluoride ion-selective electrode following Chinese standard WS/T 89-2015. A brief description was provided, but no QC methods were reported. The study authors do not account for urinary dilution in the spot samples.
    - *Direction/magnitude of effect*: Not accounting for dilution could cause there to be some exposure misclassification. The direction and magnitude would depend on where the differences occurred.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using acceptable methods that provide individual levels of exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* IQ was measured using the Combined Raven's Test, which is an appropriate test for the study population (++ for methods). Blinding was not mentioned; however, the outcome assessors would not likely have knowledge of the child's urinary fluoride. Subjects appear to have been recruited based on iodine levels and it is, therefore, unlikely that there would be any knowledge of potential fluoride exposure. Correspondence with the study authors for the Cui *et al.* (2018) study also indicated that the outcome assessors would have been blind.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes in the abstract, introduction, and methods are reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* The IQ scores are stated to be normally distributed, but there is no evidence that this was in fact tested. A t-test or one-way ANOVA was used to make comparisons between IQ and fluoride. The primary focus of the study was to evaluate associations between thyroid hormones or dopamine levels on IQ (not between fluoride and IQ). It should also be noted that regardless of the analysis conducted, there is no adjustment for school and no accounting for the clustering of children from the same school.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk if bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements, but the

study is limited by the cross-sectional study design, lack of accounting for urine dilution, and by not addressing age as a potential confounder.

Ding *et al.* (2011)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Elementary school children aged 7–14 years old
- ***Study area:*** Hulunbuir City, Inner Mongolia, China
- ***Sample size:*** 331 school children
- ***Data relevant to the review:*** IQ mean difference based on 10 categories of urine fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant association between urinary fluoride and IQ score (each increase in urinary fluoride of 1 mg/L was associated with an IQ score 0.59 points lower; 95% CI: −1.09, −0.08).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study randomly selected 340 7–14-year-olds from four nearby elementary schools in Hulunbuir. Authors stated that the four elementary schools appeared to be very similar in teaching quality. The study authors noted that they followed the principles of matching social and natural factors like economic situation, educational standards, and geological environments as much as possible; however, how this was done is unclear and no table of study subject characteristics by group was provided.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited within the same time frame using the same methods, with no evidence of differences in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* It was noted that none of the four sites had other potential neurotoxins including arsenic in their drinking water. While they did not provide the specifics, they did provide a reference. In addition, iodine deficiency was noted as not being issue in any of the four areas. Age was the only confounder adjusted in the model. While dental fluorosis severity by % female was reported, not enough data were provided to determine if it was a confounder that should have been considered in the regression. The study authors note that future studies will include covariates such as parents' educational attainment, mother's age at delivery, and household income.
  - *Potentially important study-specific confounders:* Gender
    - *Direction/magnitude of effect*: There is not enough information to determine if there is an effect from gender. There were some differences in dental fluorosis

level by gender, but it is unclear how this might impact the results or if the distribution of gender differed by age.

- *Basis for rating:* Probably high risk of bias based on indirect evidence that there were differences in gender that were not considered in the study design or analyses.

- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Data were relatively complete (i.e., <5% loss). Of the 340 subjects selected for inclusion, 5 were excluded because they lived in the area for less than a year with an additional 4 not consenting to participate.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exclusion of subjects from analysis was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Spot urine samples were collected and measured using China CDC standards. All samples were analyzed twice using a fluoride ion-selective electrode. Recovery rates were specified as 95–105% with an LOD of 0.05 mg/L. Water samples were collected from small-scale central water supply systems and tube wells with handy pumps and were processed using standard methods similar to the urine samples. Quality assurance validation was reported. A blind professional examiner evaluated the children for dental fluorosis using the Dean's Index. All urine and water samples were above the LOD. Urine levels were the primary exposure measure used in the analysis. The study authors did not account for urinary dilution in the spot samples. The mean urine fluoride concentration was correlated with the dental fluorosis levels.
    - *Direction/magnitude of effect***:** Spot urine samples that did not account for dilution could have exposure misclassification. The misclassification is likely non-differential and potential direction of bias is unknown.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measure exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* IQ was determined using the Combined Raven's Test-The Rural in China (CRT-RC3) (++ for methods). Although blinding was not reported, it is unlikely that the IQ assessors had knowledge of the children's urine levels or even of the water levels from the four sites as these were sent to a separate lab for testing (+ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient details.

  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were reasonable (ANOVA and multiple linear regression), but consideration of homogeneity of variance was not reported. The NASEM (2020) review pointed out a potential concern for the lack of accounting for clustering at the school-level since children were selected from four elementary schools. However, as pointed out in the ***Selection*** domain, the authors stated that they followed the principles of matching social and natural factors like economic situation, educational standards, and geological environments as much as possible and that the four elementary schools appeared to be very similar in teaching quality.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and that there were no other potential threats to risk of bias.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements, but the study is limited by the cross-sectional study design, lack of accounting for urine dilution, and by not addressing gender as a potential confounder.

Green *et al.* (2019)

Study Details:

- ***Study design*:** Prospective cohort
- ***Population*:** Maternal-Infant Research on Environmental Chemicals (MIREC) participants (pregnant mothers and their children aged 3–4 years)
- ***Study area:*** 10 cities, Canada
- ***Sample size:*** 512 mother−child pairs (238 from non-fluoridated areas, 162 from fluoridated areas; 264 females, 248 males)
- ***Data relevant to the review:*** Adjusted linear regression models evaluating associations between IQ in both genders together and separate with maternal urinary fluoride across all three trimesters or estimated maternal fluoride intake.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower full-scale IQ with 1-mg/L increases in maternal urinary fluoride in boys (adjusted β = −4.49), but not girls (adjusted β = 2.40) and not in sexes combine (adjusted β = −1.95); significantly lower full-scale IQ with 1-mg increases in maternal intake in sexes combined (adjusted β = −3.66 [no sex interaction]); significantly lower full-scale IQ with 1-mg/L increases in drinking water fluoride in sexes combined (adjusted β = −5.29 [no sex interaction]).

Risk of Bias:

- **Author contacts:**

  - Authors were contacted in June 2019 for additional information for the risk if bias evaluation.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Pregnant women were recruited from the same population, during the same timeframe, and using the same methods as the MIREC program. Methods were reported in detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposed groups were similar and were recruited with the same methods during the same time frame.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study considered several possible covariates including maternal age, pre-pregnancy BMI, marriage status, birth country, race, maternal education, employment, income, HOME score, smoking during pregnancy, secondhand smoke in the home, alcohol consumption during pregnancy, parity, child's gender, child's age at testing, gestational age, birth weight, time of void, and time since last void. The study also conducted secondary analyses to test for lead, mercury, arsenic, and PFOA. There is no indication of any other potential co-exposures in this study population. Iodine deficiency or excess could not be assessed but is not expected to differentially occur. The study was not able to assess parental IQ or mental health disorders. Methods used to obtain the information included questionnaires and laboratory tests.
  - *Potentially important study-specific confounders:* All key confounders were addressed.
    - *Direction/magnitude of effect:* Not applicable.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and reliable and direct evidence that key confounders including potential co-exposures were addressed.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of the 610 recruited children, 601 (98.5%) completed testing. Of the 601 mother-child pairs, 512 (85.2%) had all three maternal urine samples and complete covariate data, and 400 (66.6%) had data available to estimate fluoride intake.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Spot urine samples from all three trimesters of pregnancy were evaluated using appropriate methods, and results were adjusted for creatinine and specific gravity. Fluoride intake was estimated based on fluoride water levels and information on consumption of tap water and other water-based beverages (e.g., tea, coffee) was obtained via questionnaire.
    - *Direction/magnitude of effect***:** There is not any specific direction or magnitude of bias expected. Urinary fluoride levels are reflective of a recent exposure.

Having measurements from all three trimesters of pregnancy provides a better representation of actual exposure than a single measurement although the potential for missed high exposure is possible. However, the possibility of the occurrence of missed high exposure would be similar in all females and would be non-differential. For the fluoride intake, exposure was based on the fluoride levels in the water at the residence. If women worked outside the home and the majority of intake occurred from areas outside the home (and were different from levels in the home), there is potential to bias toward the null.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The Wechsler Preschool and Primary Scale of Intelligence was normalized for ages 2.5–<4.0 and child sex using the U.S population-based norms. Blinding was not reported, but it is unlikely that the outcome assessors had knowledge of the maternal fluoride level or were aware if the city had fluoridated water.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes were reported.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Linear regression was performed. Regression diagnostics were used to test assumptions for linearity, normality, and homogeneity. There were two potential influential observations (based on Cook's distance), and sensitivity analyses re-estimated the models without these two variables. Further sensitivity analyses were also conducted. Although city was accounted for as a covariate in the regression models, the city effect should have been a random effect rather than a fixed effect to account for potential clustering of results within each city. Although the analysis used individual-level exposure rather than city-level exposure, if the exposure levels within a city are highly correlated (which might be expected given that some cities were fully on fluoridated water and others were not), the fixed-effect model could still produce biased estimates. However, correspondence with the study authors indicated that a supplemental analysis using a random effects multi-level model showed similar results to the main model.
    - *Other potential concerns:* None identified.

  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, prospective cohort design, and addressing potential key confounders.

Li *et al.* (2004) [translated in Li *et al.* 2008a]

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Full term, normal neonates 24–72 hours old from healthy mothers
- ***Study area:*** Zhaozhou County, Heilongiang Province, China
- ***Sample size:*** 91 neonates (46 males and 45 females)
- ***Data relevant to the review:*** Comparison of neurobehavioral capacity between children in the high-fluoride area compared to the control area.
- ***Reported association with fluoride exposure:*** Yes: Significant differences in neurobehavioral assessment total scores between high-fluoride (36.48 ± 1.09) and control groups (38.28 ± 1.10); significant differences in total neurobehavioral capacity scores as measured by non-biological visual orientation reaction and biological visual and auditory orientation reaction between the two groups (11.34 ± 0.56 in controls compared to 10.05 ± 0.94 in high-fluoride group).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* There is indirect evidence that the exposure groups were similar. They were recruited during the same time frame using the same methods. From 2002 to 2003, 273 neonates were born in a hospital in Zhaozhou County, China. Ninety-one of 273 full-term neonates (46 males, 45 females) were randomly selected. Mothers ranged in age from 20 to 31 years, met multiple health criteria, and had not changed residence during pregnancy. Authors report that the two study groups are located in the same area with similar climate, living habits, economic and nutritional conditions, and cultural backgrounds, but do not provide these data in the manuscript. There is no statistically significant difference in the mode of delivery, birth weight, infant length, or sex. Subjects were separated into exposure groups after random selection.
  - *Basis for Rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited within the same time frame using the same methods with no evidence of differences in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* No confounders were specifically controlled in the analysis. The study authors note similarities in characteristics in the two populations (i.e., living habits, economic and nutritional conditions, and cultural backgrounds), but do not provide these data nor do they indicate what specific characteristics were considered. There were no significant differences in infant gender, birth method, gestational age, or infant weight and length. All tests were conducted when children were 1–3 days old. No potential co-exposures were discussed. Although arsenic is considered a potential issue in China, water quality maps indicate that there is a 25–50% probability that the drinking water in that area exceeds the WHO guideline for arsenic of 10 µg/L.
  - *Potentially important study-specific confounders:* Key confounders, including child's age, child's gender, and measures of socioeconomic status (SES), were similar between exposure groups; however, arsenic was not taken into account. Arsenic often occurs in the drinking water along with fluoride in some Chinese populations; however, based on water quality maps, arsenic does not appear to be an issue in Zhaozhou County of the Heilongjiang Province. Iodine deficiencies are not mentioned.
    - *Direction/magnitude of effect*: The presence of arsenic would potentially bias away from the null if it were present with fluoride. Deficiencies in iodine would potentially bias away from the null if it were present in areas of higher fluoride, but toward the null if it were present in areas of lower fluoride.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the key confounders are taken into account, co-exposure to arsenic is likely not an issue in this area, and methods used for collecting the information are valid and reliable.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Although authors did not discuss why they only randomly selected 91 of the 273 neonates available, results were available for all 91 subjects.
  - *Basis for rating:* Definitely low risk of bias based on results being available for all subjects.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were split into control and high-fluoride groups based on fluoride levels in their places of residence. Although the levels were provided (1.7–6.0 mg/L for the high-fluoride group compared to 0.5–1.0 mg/L for the control group), it was not reported how or when these levels were measured. Urine was collected when women were hospitalized, but before labor began. Urine samples were sent to a specific lab for measurement using fluoride ion-selective electrode. It was noted that this procedure strictly followed the internal controls of the laboratory indicating quality control. Level of detection (LOD) was not provided. Urinary fluoride levels were significantly higher in the high-fluoride mothers (3.58 ±1.47 mg/L) compared to the control-group mothers (1.74±0.96 mg/L). There was indirect evidence that exposure was consistently assessed using well-established methods that directly measure exposure. Although results were mainly based on exposure area, they were supported by urine data making exposure misclassification less of a concern.

        - *Direction/magnitude of effect*: There is high variability in both water fluoride and urine fluoride in the subjects from the high-exposure area. Although there is no overlap in the water fluoride levels in the exposure areas, there is some overlap in the urine concentrations in the mothers from the two areas. This may reflect the single measurement and pose no specific bias, or it could indicate that some mothers in the high-fluoride area have lower fluoride exposure, which could bias the results toward the null.
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measure exposure.
- ***Outcome:***
    - *Rating:* Probably low risk of bias (+)
    - *Summary:* A standard neonatal behavioral neurological assessment method was carried out by professionals in the pediatric department working in neonatal section trained specifically for these programs and passing the training exams. (+ for methods). The examinations were carried out 1 to 3 days after delivery. Because urine samples were collected on the day of delivery and sent to a separate laboratory, it is likely that the outcome assessors were blind. Although the subjects were separated by fluoride exposure area, it is not likely that the professionals were aware of the exposure as the tests were conducted in the hospital (+ for blinding).
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was assessed blindly using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
    - *Rating:* Probably low risk of bias (+)
    - *Summary:* The study authors reported numerous endpoints in sufficient detail; however, because they did not provide a list of endpoints tested there is no direct evidence that all were reported.
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that all the study's measured outcomes were reported.
- ***Other potential threats:***
    - *Rating:* Probably low risk of bias (+)
    - *Summary:*
        - *Statistical analyses:* Statistical analyses were reasonable (t-test), but consideration of homogeneity of variance was not reported. This was a translated study.
        - *Other potential concerns:* It should be noted that, although the study states that subjects were randomly selected, it is unclear why only 91 subjects were included and if they were randomly selected to obtain equal groups in the high-fluoride and control groups.
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that statistical analyses were appropriate and that there were no other potential threats to risk of bias.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in the confounding, exposure, and outcome risk-of-bias domains. Study strengths include individual fluoride measurements to support the differences in the two areas. Tests were noted

to be conducted at the hospital providing indirect evidence that blinding was not a concern during the outcome evaluation. Although there was some potential for bias due to the lack of accounting for arsenic or iodine deficiencies, co-exposure to arsenic is likely not a major concern according to groundwater quality maps.

Riddell *et al.* (2019)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Canadian Health Measures Survey (cycles 2 and 3) participants (children aged 6−17 years)
- ***Study area:*** general population, Canada
- ***Sample size:*** 3,745 children
- ***Data relevant to the review:*** Adjusted odds ratios for ADHD and attention symptoms per 1 unit increase in urinary fluoride, by water fluoride in the tap water, or community fluoridation status.
- ***Reported association with fluoride exposure:*** Yes: Significantly increased odds of ADHD diagnosis (adjusted OR = 6.10; 95% CI: 1.60, 22.8) or hyperactivity/inattentive symptoms (adjusted beta = 0.31; 95% CI: 0.04, 0.58) per 1-mg/L increase in tap water fluoride. Also, a significant association between ADHD diagnosis (adjusted OR = 1.21; 95% CI: 1.03, 1.42) or hyperactivity/inattentive symptoms (adjusted beta = 0.11; 95% CI: 0.02, 0.58) and community water fluoridation status. No significant associations with urinary fluoride levels.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Subjects were part of Cycles 2 and 3 of the Canadian Health Measures Survey. This is a nationally representative sample of residents living in 10 provinces. Specific inclusion criteria were provided. This study was restricted to children 6–17 years of age with different fluoride measurements that consisted of three participant samples. One of the samples was only available in Cycle 3.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposed groups were similar and were recruited with the same methods during the same time frame.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Covariates included in all models included child's age at test, child's sex, ethnicity, BMI, parents' education, total household income, exposure to cigarette smoke inside the home, and log-transformed concurrent blood lead levels. Confounders such as parental behavioral and mental health disorders, quantity and quality of caregiving environment, and co-exposure to arsenic were not discussed. Rationale for selection of covariates was based on relationship to ADHD diagnosis and to fluoride metabolism based on literature review and consultation with an ADHD expert. There is no

information of the source if data for covariates, but this is likely the questionnaires from the Canadian Health Measures Survey, which are considered standardized and validated.

- o *Potentially important study-specific confounders:* All key confounders were considered in this study.
  - *Direction/magnitude of effect***:** Not applicable.
- o *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key confounders are taken into account, co-exposure to arsenic is likely not an issue in this area, and methods used for collecting the information were valid and reliable.

- ***Attrition:***
  - o *Rating:* Probably low risk of bias (+)
  - o *Summary:* There is no information indicating that there were any data excluded due to missing covariates. All exclusions of children were described and reasonable (i.e., drinking bottled water when considered city fluoridation as a measure of fluoride exposure). Outliers were stated to be excluded, but methods for determining this were provided and it was noted that the outliers were 0.27% of the values.
  - o *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - o *Rating:* Probably low risk of bias (+)
  - o *Summary:* **Urinary Fluoride:** Spot urine samples were collected under normal non-fasting conditions and analyzed using an Orion pH meter with a fluoride ion-selective electrode after being diluted with an ionic adjustment buffer. Analysis was performed at the Human Toxicology Laboratory of the Institut National de Sante Publique du Quebec. The precision and accuracy of the fluoride analyses, including quality control and quality assurance, were described by Health Canada (2015). The limits of detection were 20 µg/L for Cycle 2 and 10 µg/L for Cycle 3 with no values below detection. Fluoride levels were adjusted for specific gravity.

    Water Fluoride in Tap water: Tap water was collected at the subjects' homes in Cycle 3 only. Samples were analyzed for fluoride concentrations using anion exchange chromatography procedure with a LOD of 0.006 mg/L. Values below the LOD were imputed with LOD/square root 2. Of the 980 samples, 150 (16%) were below detection. **Chlorinate Water Fluoride status:** This was determined by viewing reports on each city's website or contacting the water treatment plant (provided in supplemental material). Children were excluded if they drank bottled water, had a well, had a home filtration system, lived in the current residence for 2 years or less, or lived in an area with mixed city fluoridation.
    - *Direction/magnitude of effect***:** There is not any specific direction or magnitude of bias expected. Urinary fluoride levels are reflective of a recent exposure, but the study authors adjusted to account for dilution. The possibility of exposure misclassification would be similar in all subjects and would be non-differential. There is less potential for exposure misclassification in regard to tap water or chlorinated water fluoride status as children who drank bottled water were

excluded and children who had a home filtration system were excluded from the chlorinated water status.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.

- ***Outcome:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:*
    Strengths and Difficulties Questionnaire (SDQ): The questionnaire was administered to youths under 18 years. Children aged 6–11 years had SDQ ratings provided by parents and guardians, but youths aged 12−17 years completed the questionnaire themselves. Tests consist of 25 items with a 3-point scale. Items were divided into five subscales: emotional problems, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior. The current study only used the hyperactivity-inattention subscale. Validation of this method was not reported (- for methods).
    ADHD: Ninety percent of youths with ADHD are diagnosed after age 6 years. For children aged 6–11 years, ADHD diagnosis was provided by parents, but youths age 12−17 years completed the questionnaire themselves. Cycle 2 asked "Do you have a learning disability?" and if yes asked to specify the type (4 options available and described). In Cycle 3, parents were asked directly whether they had ADHD, and children 12 years and older were asked if they had a physician diagnosis of ADHD and, if so, what subtype. (- for methods because different methods were used and only the children 12 years and older in cycle 3 were asked specifically about doctor diagnosis). Both were measured in both cycles. Blinding is not likely an issue as subjects would not have knowledge of the urine or tap water fluoride levels. However, they would likely have knowledge of the city.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the outcome was assessed using insensitive methods that varied based subject age.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods sections were reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Logistic regression was used for ADHD results. Box-Tidewell tests were used to check the linearity of the relationship with the continuous predictors. Linear regression was used for the SDQ scores using Huber-White standard errors. All regressions were tested for interactions between age and fluoride and sex and fluoride. Sensitivity analyses were conducted to test the different cycles.
    - *Other potential concerns:* None identified.

  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include individual exposure measurements and the addressing of potential key confounders but was limited by the cross-sectional study design and insensitive outcome measures.

Rocha-Amador *et al.* (2007)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 6–10 years
- ***Study area:*** Moctezuma (low fluoride, low arsenic) and Salitral (high fluoride, high arsenic) of San Luis Potosí State and 5 de Febrero (high fluoride, high arsenic) of Durango State, Mexico
- ***Sample size:*** 132 children
- ***Data relevant to the review:*** Associations between full-scale IQ, performance IQ, verbal IQ and child's urine or water fluoride levels.
- ***Reported association with fluoride exposure:*** Yes: Significant associations between fluoride and IQ scores (full-scale IQ adjusted βs of –10.2 with water and −16.9 with urine; CIs not reported); arsenic also present, but the effect was smaller (full-scale IQ adjusted βs of –6.15 with water and –5.72 with urine; CIs not reported).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* All children in 1st through 3rd grades in three rural areas in Mexico (n = 480) were screened for study eligibility including age, time at residence, and address. Authors report that the three selected communities were similar in population and general demographic characteristics. Children who had lived in the area since birth and were 6–10 years old were eligible to participate (n = 308). Of the 308 children, 155 were randomly selected and the response rate was 85%, but participation was not reported by area. It was noted, however, that no significant differences in age, gender, or time of residence were observed between participants and non-participants. Timeframe for selection was not mentioned but appears to be similar. Sociodemographic characteristics of subjects was provided in Table 1 of the study. There was a significant difference in SES and transferrin saturation, but these were taken into account in the analysis.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the populations were similar and differences were noted and addressed in the analysis.
- ***Confounding:***

  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study design or analysis accounted for child's age, sex, SES, transferrin saturation, weight, height, blood lead levels, and mother's education. Arsenic levels were highly correlated with fluoride levels and it was stated that each was tested alone, and arsenic was found to have less of an effect. The authors noted in the methods that they tested for an interaction between arsenic and fluoride. Smoking was not addressed and methods for measuring many of the confounders were not reported.
  - *Potentially important study-specific confounders:* Arsenic
    - *Direction/magnitude of effect*: Presence of arsenic, which also demonstrated an association, would bias away from the null.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and reliable and direct evidence that key confounders were addressed.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of 155 children randomly selected for study participation, 85% responded to enroll. According to the authors, there were no significant differences in age, gender, or time of residence between responders and non-responders. However, no data are provided to support this, and no breakdown of responders/non-responders by region is provided. Data were provided for the 132 children agreeing to participate.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Urine samples were collected on the same day as psychological evaluations and were analyzed for fluoride according to NIOSH Method 8308 (Fluoride in Urine). For QC, a reference standard was also used (NIST SRM 2671a). Urine samples were also analyzed for arsenic by using the Atomic Absorption Spectrophotometer with hydride system (Perkin-Elmer, model AAnalyst 100, Wellesley, United States) and used a reference standard for QC. Levels were adjusted for urinary creatinine levels to account for dilution in the spot samples. Tap water samples were collected from each child's home on the day of biological monitoring. Fluoride was measured with a sensitive, specific ion electrode. Detailed methods are provided including internal quality controls. It was noted that in the high fluoride group it was common to drink bottled water low in fluoride and to only use the tap water for cooking; therefore, urine was considered the most appropriate measure of exposure. Only children who had lived at the same residence since birth were included.
    - *Direction/magnitude of effect*: Not applicable.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* Neuropsychological profiles were assessed through the WISC-RM (revised for Mexico). This is a well-established test appropriately adjusted for the study population. However, no additional validation is provided (+ for methods). The study report stated that the test assessors were masked to both arsenic and fluoride water levels (++ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* It was reported that an interaction between fluoride and arsenic was measured, but it was only noted in the discussion that the study design precluded testing statistical interaction between fluoride and arsenic. This provides indirect evidence of selective reporting.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that there was selective reporting.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
  - *Statistical analyses:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements and outcomes blindly assessed, but it is limited by the cross-sectional study design and not being able to completely rule out the influence of arsenic in the results.

Rocha-Amador *et al.* (2009)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 6–11 years
- ***Study area:*** Durango, Mexico
- ***Sample size:*** 80 children
- ***Data relevant to the review:*** Associations between visuospatial organization and visual memory (using the Rey-Osterrieth Complex Figure Test, children's version) and urinary fluoride levels in the children.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between urinary fluoride and visuospatial organization (r = −0.29) and visual memory (r = −0.27) scores. No significant correlations with arsenic.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were from the same population and were recruited during the same time frame using the same methods. Although this study compared three sites with antecedents of environmental pollution to mixtures of either F–As, Pb–As, or DDT–PCBs, authors evaluated each contaminant separately. The only area of interest is the area with F and As contamination. The area in Durango state (5 de Febrero) where drinking water is polluted naturally with F and As at levels exceeding 6 and 19 times, respectively, the World Health Organization (WHO) limits (WHO 2008). Children attending public schools were screened through personal interviews for study eligibility. Inclusion criteria were children between 6 and 11 years old, living in the study area since birth, and whose parents signed the agreement to participate. Children with a neurological disease diagnosed by a physician and reported by the mother were excluded from the study. The final sample for the F–As was 80. Participation rates were not reported. Selected demographic characteristics are presented in Table 1 of the study.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the populations were similar and recruited during the same time frame using the same methods.
- ***Confounding:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* Confounding factors in children tested in the analysis included blood lead (PbB), age, gender, and height-for-age z-scores; only age had significant associations and was included in the final analysis. Arsenic was also assessed and analyzed separately from fluoride. Arsenic in urine was analyzed by atomic absorption spectrophotometer coupled to a hydride system (Perkin-Elmer model AAnalyst 100). Although the model did not adjust for arsenic, arsenic in the F–As group was not associated with either endpoint. PbB was analyzed with a Perkin-Elmer 3110 atomic absorption spectrophotometer using a graphite furnace. Authors note that the mean blood lead level in the F–As study area was 5.2 µg/dL and 8% of the children had values above the reference value of 10 µg/dL. PbB was stated not to affect results and was not included in the final analysis. Other confounding data were obtained during the study interview. Father's education was provided and, in the F–As group, was stated to range from 0–16 years, but this was not considered. Maternal education, smoking, and SES were also not considered. The authors provide an SES score of 5.9 ± 1.4 for the 5 de Febrero region (the fluoride region). It is not clear if this would vary by fluoride or arsenic levels.
  - *Potentially important study-specific confounders:* SES.
    - *Direction/magnitude of effect***:** There are insufficient data to determine the magnitude or direction of effect. If there is an association between fluoride exposure and SES, the direction of effect would depend on the association.

  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the SES was not accounted for in the study design or analysis and may have varied by fluoride levels.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Data are complete. All 80 participants stated to be the final sample for the site of interest (F–As) were included in all analyses.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that there was no attrition.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Fluoride in urine (FU) was analyzed according to method 8308 (''fluoride in urine'') from the National Institute of Occupational Safety and Health (NIOSH 1984) with a sensitive specific ion electrode. As a quality control check, reference standard ''fluoride in freeze dried urine'' (NIST SRM 2671a) was analyzed. The accuracy was 97.0 +/- 6.0%. Levels of FU and AsU were adjusted for urinary creatinine, which was analyzed by a colorimetric method (Bayer Diagnostic Kit, Sera-Pak1 Plus). However, details on the collection methods were not reported.
    - *Direction/magnitude of effect***:** Spot urine samples in a small sample size (i.e., 80 children) may have some exposure misclassification. Adjusting for dilution reduces the potential for misclassification based on differences in dilution. Exposure misclassification would be non-differential.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* IQ is assessed through the Rey-Osterrieth Complex Figure Test (ROCF). This is a less well-established method, although the authors provide citations suggesting it has been validated and standardized for the Mexican population (+ for methods). According to the study report, the neuropsychologist who administered the test was blinded to all exposure types and levels. (++ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study.
    - *Other potential concerns:* None identified.

  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and outcomes blindly assessed, but it is limited by the cross-sectional study design, lack of addressing SES in the study population, co-exposure with arsenic, and use of spot samples in a small population.

Saxena *et al.* (2012)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 12 years
- ***Study area:*** Madhya Pradesh, India
- ***Sample size:*** 170 children
- ***Data relevant to the review:*** Mean IQ grade (not standard scores) by water fluoride quartiles or continuous or by continuous urinary fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between water (r = 0.534; p = 0.000) and urinary (r = 0.542; p = 0.000) fluoride levels and IQ score.

Risk of Bias:

- **Author contacts:**
  - Authors were contacted in August of 2017 to obtain additional information for risk-of-bias evaluation.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* There was indirect evidence that subjects were similar and were recruited using the same methods during the same time frame. The study participants were selected from a stratified cluster of geographic areas based on fluoride concentration in groundwater. According to the authors, the selected villages were similar in population and demographic characteristics. Data are provided to show the breakdown in SES, parental education, height/age, and weight/height and no significant differences were noted. Participation was stated to be voluntary, but participation rates were not provided. It is unclear if the 170 subjects were selected with 100% participation or if the 170 subjects were all that were asked to participate, but it appears that all subjects participated. Timing of the recruitment was not provided but is assumed to occur during the same time frame.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that subjects were similar and recruited using the same methods during the same time frame.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* There was indirect evidence that key confounders including potential co-exposures were addressed using reasonable methods. A questionnaire, completed with the assistance of parents, was used to collect information on child characteristics (age, sex, height, weight), residential history, medical history (including illness affecting nervous system and head trauma), educational level of the head of the family (in years), and SES of the family. The SES was recorded according to the Pareek and Trivedi classification. The nutritional status of the children was calculated using the Waterlow's classification, which defines two groups for malnutrition using height for age ratio (chronic condition) and weight for height ratio (acute condition). Within both groups, it categorizes the malnutrition as normal, mildly impaired, moderately impaired, or severely impaired. Urinary lead and arsenic were analyzed using the atomic absorption spectrophotometer (Perkin-Elmer, Wellesley, United States). Urinary iodine was measured using the Dunn method. Authors do not report which covariates were included in the multivariate regression models; however, there was no difference in reported demographic characteristics. All subjects were the same age, and there was no difference in iodine, lead, or arsenic between the groups. Mean urinary arsenic levels did increase with increasing fluoride even though there was no significant difference by group.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and reliable and that key confounders including potential co-exposures were addressed.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Results were provided for all 170 children stated to be included in the study.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence of no attrition.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* A sample of 200 mL of drinking water was collected at each child's home. The fluoride levels were analyzed by a fluoride ion-selective electrode, Orion 9609BN (Thermo Fisher Scientific Inc., West Palm Beach, United States). Each subject was also asked to collect a sample of their first morning urine. The fluoride content in the urine was determined using a fluoride ion-selective electrode, Orion 9609BN (Thermo Fisher Scientific Inc., West Palm Beach, United States). QA/QC and LOD were not reported and urinary dilution was not assessed. Although only current levels were measured, children who had changed water source since birth were excluded.
    - *Direction/magnitude of effect***:** Spot urine samples that did not account for dilution could have exposure misclassification. The misclassification is likely non-differential and not likely to bias in any specific direction. Children who had changed water since birth were excluded, but it was not specifically noted that the fluoride in the water source was stable over the years.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Intelligence is assessed using the Raven's Standard Progressive Matrices and categorized into five grade levels. Although it was not noted that the test was validated to the study population, the test is visual and would be applicable to most populations (+ for methods). There is no mention of blinding by test administrators or evaluators and the exposure groups come from different geographic areas. It was also not reported who measured the levels of fluoride from the home or urine samples. Correspondence with the study authors indicated that the outcome assessors were blind to the children's fluoride status (++ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were reasonable (ANOVA), but consideration of homogeneity of variance was not reported.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate, and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements and the addressing of potential key confounders, but it was limited by the cross-sectional study design and lack of addressing dilution in the urine samples.

Seraj *et al.* (2012)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 6–11 years
- ***Study area:*** five villages, Makoo, Iran
- ***Sample size:*** 293 children

- ***Data relevant to the review:*** IQ (mean and distribution) assessed by Raven's Colored Progressive Matrices and presented by fluoride area, beta was also provided for water fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between water fluoride and IQ score (adjusted β = −3.865; CIs not reported); significantly higher IQ score in normal area (97.77 ± 18.91) compared with medium (89.03 ± 12.99) and high (88.58 ± 16.01) areas.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were selected from five villages in Makoo. The villages were stated to all be rural with similar general demographic and geographic characteristics and were comparable in terms of SES and parental occupations. Children were 6–11 years old. Age, gender, and education were taken into account in the analysis. No other characteristics were provided or discussed. Participation rates were not reported. There is indirect evidence that the populations were similar, and some possible differences were addressed.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that subjects were similar and recruited using the same methods during the same time frame.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Age, gender, dental fluorosis intensity, and educational levels (child's and parents') were evaluated as potential confounders. Other potential confounders such as smoking were not discussed. Information was obtained from a detailed questionnaire. Lead was measured, but only found in low levels in the drinking water throughout the study regions. Iodine in the water was also stated to be measured and residents were receiving iodine-enriched salt. Arsenic was not addressed, but there is no evidence that arsenic levels would vary across villages in this area.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and that key confounders including potential co-exposures were addressed or were not likely to be an issue in the study area.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Attrition was low if it occurred. It was noted that 293 out of 314 children living in the villages were recruited. It is not clear if 21 children were excluded based on exclusion criteria or if they refused to participate; however, this accounts for less than 10% of the population and results were available for all 293 subjects.

  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* Exposure was primarily based on area of residence. Fluoride in the groundwater was analyzed by the SPADNS (Sulfophenylazo dihydroxynaphthalene-disulfonate) method, utilizing 4000 UV-Vis spectrophotometer (Hach Company, Germany) in the environmental health engineering laboratory of the Public Health School of Tehran University of Medical Sciences. Specific details were not provided on methods of collection, samples locations, or if these locations represented the primary sources of drinking water for the subjects. Villages were categorized into normal (0.5–1 ppm), moderate (3.1±0.9 ppm), and high (5.2±1.1 ppm) fluoride based on the mean fluoride content of all seasons presumably for the stated 12-year time period. Subjects were stated to be long-life residents of the village. Dental fluorosis was also measured and increased in severity with fluoride levels; however, all areas had some degree of dental fluorosis. Although authors used an average fluoride level in varying seasons over presumably 12 years, they used a less-established method without reporting reliability or validity, nor did they provide data to indicate that the mean was truly representative of the fluoride levels over time and throughout the village. Although dental fluorosis severity increased with increasing fluoride levels, the data could also indicate potential exposure misclassification.
    - *Direction/magnitude of effect***:** The presence of dental fluorosis in all groups indicates that there may have been different exposure in some children at a younger age. Although there were only about 20 children in the "normal" fluoride group with very mild to mild dental fluorosis, this could bias the results toward the null because those children may have experienced a higher level of fluoride at some point. The other two fluoride groups were exposed to fluoride levels that likely exceeded those in the "normal" fluoride group.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that exposure was assessed using insensitive methods.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Intelligence was evaluated using the Raven's Color Progressive Matrices. This is a well-established method. Although the study authors did not provide data to indicate that the methods were valid in this study population, the test is designed to be culturally diverse. (+ for methods). The study report stated that test administrators were blinded. (++ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that outcomes were blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported. However, because they did not report the method for obtaining the betas in Table 4 of the study, it is not clear if these were adjusted or unadjusted betas.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that all the study's measured outcomes were reported, but the results were not sufficiently reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical comparisons between groups were reasonable (ANOVA), but consideration of homogeneity of variance was not reported. In addition, the methods for obtaining the betas were not reported.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and outcome. Study strengths include addressing potential key confounders, but it was limited by the cross-sectional study design and the group-level exposure data.

Soto-Barreras *et al.* (2019)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 9–10 years
- ***Study area:*** Chihuahua, Mexico
- ***Sample size:*** 161 children
- ***Data relevant to the review:*** Water fluoride, urinary fluoride, exposure dose, and dental fluorosis index by IQ grade.
- ***Reported association with fluoride exposure:*** No: No significant associations between fluoride exposure and IQ grades.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were selected using a multistage cluster sampling. During the first stage, 13 public elementary schools were randomly selected from a pool of 73 using a cluster sample design. Secondly, only fourth grade students were included. Authors stated that they wanted to keep the same grade level, but they were not specific as to why fourth graders were selected as opposed to any other grade. Lastly, only children whose parents or guardians attended and responded to the survey were included. There is no information provided on how the 13 schools selected may be similar or different from the 60 schools not selected. There is no information provided on the number of

children in the fourth grade to know participant rates. It was only noted that 245 children were examined, but 161 were included after the exclusion rules were applied. Inclusion and exclusion criteria are presented. Reasons for exclusion do not appear to be related to exposure or outcome. Characteristics of participants and non-participants are not compared; however, characteristics of the 161 included children were provided and any differences were taken into account in the analysis.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposed groups were similar and were recruited using similar methods during the same time frame.

- ***Confounding:***
  - *Rating:* Probably high risk of bias (-)
  - *Summary:* No confounders was considered when evaluating fluoride associations with intelligence; they were only applied when evaluating fluoride levels and dental caries. Based on Table 4 of the study, there was no significant association between IQ grade and child's age, sex, parental education, or SES status. No other information was reported or considered. There is no information on potential co-exposures. Based on water quality maps, the arsenic prediction indicates a greater than 50% probability of exceeding the WHO guidelines for arsenic of 10 µg/L in areas of Chihuahua, Mexico.
  - *Potentially important study-specific confounders:* Arsenic.
    - *Direction/magnitude of effect***:** The direction and magnitude of effects is unknown. There is potential for arsenic to occur in the study area, but it is not known how it relates to fluoride exposure. If they occur together in the water, it will bias away from the null; however, if they occurred in different areas, there is potential to bias toward the null.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that there is potential for exposure to arsenic that was not sufficiently addressed.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* A total of 161 of 245 children were included in the study. Exclusion criteria are presented and are unrelated to outcome or exposure. For the 161 children, there are no missing outcome data.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+);Probably high risk of bias (-)
  - *Summary:* **Urinary Fluoride (probably low risk of bias):** First morning void urine samples were collected based on NIOSH methods. Water samples were also stated to be collected, but it does not appear that methods followed any particular standard, and there is no indication that subjects were provided with collection containers. Analysis was based on a calibration curve using fluoride ion selective electrode. QC methods were mentioned. Based on results, there were values below detection limits, but LODs or % below LOD were not reported.

Daily fluoride exposure (probably high risk of bias): Daily fluoride exposure was based on the water fluoride level, drinking water consumption (based on parental report of how many glasses of water consumed), and body weight.

- *Direction/magnitude of effect***:** Spot urine samples that did not account for dilution could have exposure misclassification. The misclassification is likely non-differential and is not likely to bias in any specific direction. Daily exposure was based partially on parental report of water consumption. The direction and magnitude of effect is unknown.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure. The daily fluoride exposure is probably high risk of bias because there is indirect evidence that the exposure was assessed using methods of unknown validity.

- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Intellectual ability was evaluated using Raven's Colored Progressive Matrices by an independent examiner. Some details were provided, but it was not stated that the tests were assessed blind; however, there is no indication that subjects were from high fluoride areas and the assessor would not have knowledge of the urine or water fluoride levels. Results for children were converted into a percentile according to age (details not provided) and overall scores were assigned an intellectual grade of I to V as described in the report.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* The main analysis was for dental caries. Although they make conclusions on fluoride and IQ, they do not use the same analytical methods for both outcomes. Table 4 of the study provides a p-value although it is not clear what the p-value represents; it is presumed to be the Kruskal Wallis p-value. It appears that a Kolmogorov-Smirnov test was used to determine variable distribution and a Kruskal Wallis test was used to compare among the groups with a Dunn's test if significant.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk if bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.

- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and outcomes blindly assessed, but is limited by the cross-sectional study design, lack of accounting for urine dilution, and by not addressing potential exposures to arsenic in the study area.

Sudhir *et al.* (2009)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 13–15 years
- ***Study area:*** Nalgonda district (Andhra Pradesh), India
- ***Sample size:*** 1,000 children
- ***Data relevant to the review:*** Mean IQ grade (not standard scores) or IQ distribution by water fluoride strata (<0.7, 0.7-1.2, 1.3-4.0, and >4.0 ppm).
- ***Reported association with fluoride exposure:*** Yes: Significantly increased number of intellectually impaired children with increasing drinking water fluoride levels.

Risk of Bias:

- **Author contacts:**
  - Authors were contacted in September of 2017 for additional information related to risk-of-bias evaluation, but no response was received.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Children were selected from the same general population during the same time frame and were then broken down into nearly equal exposure groups. A cross-sectional study was conducted among 13–15-year-old school children of Nalgonda district, Andhra Pradesh between August and October 2006. Data were collected from the school children who were life-long residents of Nalgonda district, Andhra Pradesh and who consumed drinking water from the same source during the first 10 years of life. A stratified random sampling technique was used. The entire geographical area of Nalgonda district was divided into four strata based on different levels of naturally occurring fluoride in the drinking water supply. Children were randomly selected from schools in the different strata. It was noted that the 1,000 selected children were equally divided among all four strata, however, each group did not have 250 children (but instead 243−267 in each group). Participation rates are not reported. Exclusion criteria included: children who had a history of brain disease and head injuries, children whose intelligence had been affected by congenital or acquired disease, children who had migrated or were not permanent residents, children with orthodontic brackets, and children with severe extrinsic stains on their teeth. Age and gender data are presented in Table 1 of the study, but this information is not presented by the different fluoride groups.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that subjects were similar and were recruited using the same methods during the same time frame.

- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Data were collected using a self-administered questionnaire and clinical examination. The self-administered questionnaire requested information on demographic data (appears to cover age and sex), permanent residential address, staple food consumed, liquids routinely consumed, and aids used for oral hygiene maintenance (fluoridated or nonfluoridated). SES was measured using the Kakkar socio-economic status scale (KSESS) with eight closed-ended questions related to parental education, family income, father's occupation, and other factors. All children were asked to fill out the form, and the answers obtained were scored using Kakkar socio-economic status scoring keys. Based on this scoring, children were divided into three groups—lower class, middle class, or upper class. Age, sex, and SES were not found to be significantly associated with IQ. Other confounders including smoking were not addressed. Co-exposures such as arsenic and lead were not addressed; however, there is no indication that lead is a co-exposure in this population and arsenic is not likely a major concern in this area based on water quality maps.
  - *Potentially important study-specific confounders:* Key confounders age, gender, and measures of SES were similar between exposure groups; however, arsenic was not taken into account. Arsenic often occurs in the drinking water along with fluoride in some Indian populations; however, based on water quality maps, this does not appear to be an issue in the Nalgonda district of Andhra Pradesh. Iodine deficiencies are not mentioned.
    - *Direction/magnitude of effect***:** The presence of arsenic would potentially bias away from the null if present with fluoride. Deficiencies in iodine would bias away from the null if present in areas of high fluoride, but toward the null if present in areas of non-high fluoride.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the key confounders are considered, co-exposure to arsenic is likely not an issue in this area, and methods used for collecting the information were valid and reliable.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Results were available for the 1,000 children selected to participate.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence of no attrition.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Children were placed into one of four strata based on the level of fluoride in drinking water. Collection of water samples was done in the districts. The placement into strata was based on fluoride levels obtained from documented records of District Rural Water Works Department. Once the children were assigned to strata, it was confirmed that the fluoride level of their drinking water was within the strata assigned. This was done using the methodology followed in National Oral Health Survey and Fluoride Mapping 2002–2003. During the initial visits to the schools, the children were interviewed regarding their history of residence and source of drinking water from birth to 10 years. The first child meeting criteria was given a bottle for water collection and

the next child was only given a bottle for collection if the water source was different than that of a previous child. Children were asked to collect the sample of water from the source that was used in the initial 10 years of their life and was collected the next day. It was not reported if all bottles were returned. The water samples collected were subjected to water fluoride analysis using an ion-specific electrode, Orion 720A fluoride meter at District Water Works, Nalgonda to confirm the fluoride levels in the water before commencement of clinical examination. LOD and QA/QC details were not reported.

- *Direction/magnitude of effect***:** There is some potential for exposure misclassification based on recall of the children on the source of water used in their first 10 years of life. The misclassification is likely non-differential and not likely to bias in any specific direction. Children who had changed water since birth were excluded, but it was not specifically noted that the fluoride in the water source was stable over the years.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.

- ***Outcome:***
  - *Rating:* Probably high risk of bias (NR)
  - *Summary:* The Raven's standard progressive matrices (1992 edition) was used to assess IQ. Exams were carried out by a single examiner. Calibration of the examiner was done before the study and in the middle of the study, but it was not clear if this applied to the IQ evaluation or only to the clinical examination. This Raven's test is a standard test and although there is no information provided to indicate that the methods were reliable and valid in the study population, this test was created to be culturally fair (+ for methods). Blinding or other methods to reduce potential bias were not reported (NR for blinding). No response was received to an e-mail request for clarification in September 2017. Overall rating for methods and blinding = NR.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the outcome was not assessed blind and could bias the results.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were appropriate and no other threats to internal validity were identified.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.

- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include verification of exposure measurements and the addressing of potential key confounders, but it was limited by the cross-sectional study design and lack of information on blinding during outcome assessment.

Till *et al.* (2020)

Study Details:

- ***Study design*:** Prospective cohort
- ***Population*:** MIREC participants (pregnant mothers and their children aged 3–4 years)
- ***Study area:*** 10 cities, Canada
- ***Sample size:*** 398 mother−child pairs (247 from non-fluoridated areas, 151 from fluoridated areas; 200 breastfed as infants, 198 formula-fed as infants)
- ***Data relevant to the review:*** Adjusted linear regression models evaluating associations between IQ with water fluoride concentration (with or without adjusting for maternal urine) in formula-fed or breast-fed infants or by fluoride intake from formula.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower performance IQ with water fluoride (adjusted βs = −9.26 formula-fed, −6.19 breastfed) and fluoride intake from formula (adjusted β = −8.76); significantly lower full-scale IQ with water fluoride in formula-fed (adjusted β = −4.40); no significant changes in full-scale IQ for water fluoride in breastfed children or fluoride intake from formula-fed children; no significant changes in verbal IQ scores with fluoride exposure.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Pregnant women were recruited between 2008 and 2011 by the MIREC program from 10 cities across Canada. Inclusion and exclusion criteria were provided. Additional details were stated to be available in Arbuckle *et al.* (2013). A total of 610 children were recruited to participate in the developmental follow-up with 601 children completing all testing. The demographic characteristics of women included in the current analyses (n = 398) were not substantially different from the original MIREC cohort (N = 1945) or the subset without complete water fluoride and covariate data (n = 203). A table of characteristics of the study population is provided. Approximately half of the children lived in nonfluoridated cities and half lived in fluoridated cities.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposed groups were similar and were recruited with the same methods during the same time frame.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)

  - *Summary:* Covariates were selected a priori that have been associated with fluoride, breast feeding, and children's intellectual ability. Final covariates included child's sex and age at testing, maternal education, maternal race, second-hand smoke in the home, and HOME score. City was considered but was excluded from the models. Confounders that were not assessed include: parental mental health, iodine deficiency/excess, parental IQ, and co-exposure to arsenic and lead. Co-exposure to arsenic is less likely an issue in this Canadian population and the lack of information is not considered to appreciably bias the results.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that key confounders were addressed and indirect evidence that the methods used to collect the information were valid and reliable and co-exposures were not an issue.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of 610 children, 601 (98.5%) in the MIREC developmental study who were ages 3–4 years completed the neurodevelopment testing. Of the 601 children who completed the neurodevelopmental testing, 591 (99%) completed the infant feeding questionnaire and 398 (67.3%) reported drinking tap water. It was noted that the demographic characteristics were not substantially different from the original MIREC cohort or the 203 subjects without complete water fluoride or covariate data.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Information on breastfeeding was obtained via questionnaire at 30–48 months. Fluoride concentration in the drinking water was assessed by daily or monthly reports provided by water treatment plants. Water reports were first linked with mothers' postal codes and the daily or weekly amounts were averaged over the first 6 months of each child's life. Additional details can be found in Till *et al.* (2018). Maternal urinary exposure was used to assess fetal fluoride exposure. Procedures can be found in Green *et al.* (2019).
    - *Direction/magnitude of effect***:** There is not any specific direction or magnitude of bias expected. Urinary fluoride levels are reflective of recent exposure. The possibility of the exposure misclassification would be similar in all subjects and would be non-differential. For the fluoride intake from formula, exposure was based on the fluoride levels in the water at the residence and the proportion of time that the infant was not exclusively breastfed. This exposure misclassification would also be non-differential.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***

  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Intelligence was tested using the Wechsler Preschool and Primary Scale of Intelligence III. This is appropriate for both the study population and age group. This is considered a gold standard test. It was not reported whether the evaluators were blind to the child's fluoride exposure status during the assessment. Although it is unlikely that the assessors had knowledge of the specific drinking water levels or maternal urine levels, there is potential that the outcome assessors had knowledge of the city the child lived in and if the city was fluoridated or non-fluoridated. Correspondence with the study authors on the outcome assessment for Green *et al.* (2019) indicated that it was unlikely that the testers had knowledge of the city's fluoridation. The same is assumed here. Specific measurements included were identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Regression diagnostics were used to test assumptions for linearity, normality, and homogeneity. There were two potential influential observations (based on Cook's distance), and sensitivity analyses re-estimated the models without these two variables. Effect modification by breastfeeding status was evaluated. Interestingly, all regression coefficients were divided by 2 to represent change in IQ per 0.5-mg/L change in fluoride. One concern is posed by the lack of accounting for city in the regression models, ideally as a random effect. The authors explored including city as a covariate in the models; however, city was not included either because it was strongly multi-collinear with water fluoride concentration (VIF > 20) (model 1, with water fluoride concentration) or because fluoride intake from formula is a function of water fluoride concentration (assessed at the city level) and was therefore deemed redundant (model 2). However, the models use city-level water fluoride concentrations (and in sensitivity analyses, adjust for maternal urinary fluoride) which warrant exploration of city as a random effect rather than a fixed effect (as would be by just having it included as a covariate). Even including individual-level maternal urinary fluoride might not fully account for lack of city, given that the subjects were from six different cities, with half of them fully on fluoridated water. Hence, even individual-level exposures are likely to be correlated at the city level. Based on a previous analysis (Green *et al.* 2019), it is unlikely that

exclusion of city from models (as a fixed or random effect) would impact the effect estimates.
        - *Other potential concerns:* None identified.
    - *Basis for rating:* Probably low risk if bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, prospective cohort design, and the addressing of potential key confounders.

Trivedi *et al.* (2012)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 12–13 years
- ***Study area:*** Kachchh, Gujarat, India
- ***Sample size:*** 84 children
- ***Data relevant to the review:*** Mean IQ scores and distribution by low and high fluoride villages.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower IQ score in the high fluoride (92.53 ± 3.13) compared to the low fluoride (97.17 ± 2.54) areas in boys and girls combined (as well as separately). Villages with higher fluoride levels had a larger percentage of subjects with IQ scores of 70–79, while the lower fluoride villages had a greater percentage of IQ scores > 109.

Risk of Bias:

- **Author contacts:**
    - Authors were contacted in September of 2017 to obtain additional information for risk-of-bias evaluation.
- ***Population selection:***
    - *Rating:* Probably low risk of bias (+)
    - *Summary:* There is insufficient information provided on the sampling methods to determine if the populations were similar. Although it was noted that samples were obtained for groundwater quality from March to May of 2011, there is no indication that the children were selected at the same time or during a similar time frame. Correspondence with the author indicates that children were selected within a week of the water collection based on random selection of a school in the village. Study participants were selected from six different villages of the Mundra region of Gujarat, India. Subjects were grouped into high and low villages based on the level of fluoride in the drinking water of those villages. The number of subjects per village were not reported, but it was noted that there were 50 children in the low fluoride group and 34 children in the high fluoride group. It is not clear if the differences in numbers were based on different participation rates or if there were fewer children in the high fluoride villages. Recruitment methods including any exclusion criteria and participation rates

were not provided. SES was stated to be low and equal based on questionnaire information, but the results were not provided. It should also be noted that only regular students (having attendance more than 80%) of standard 6th and 7th grades were selected, but it was not noted if attendance varied by village. Correspondence with the study author indicated that there was an average of 20 students per class with an average of 40 students per village. It appears that keeping the requirement for 80% attendance was a limiting factor that caused different numbers of children by area; however, this was applied similarly to both groups.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that subjects were similar and recruited using the same methods during the same time frame.

- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Children were stated to be students of the 6th and 7th standard grades. Age was not addressed, but the children would all be of similar age based on the grades included. Results were reported for males and females separately as well as combined. SES and iodine consumption were stated to be analyzed via a questionnaire and were standardized on the basis of the 2011 census of India. Although it was noted in the abstract that the SES was equal (no data provided), the study report did not mention the iodine results. Although the study authors did not address arsenic or lead, they did provide physicochemical analyses for the water samples from the six different villages. Information on arsenic in the water is not provided, but based on water quality maps, arsenic is not expected to be a major concern in this study area.
  - *Potentially important study-specific confounders:* Key confounders age, gender, and measures of SES were similar between exposure groups; however, arsenic was not taken into account. Arsenic often occurs in the drinking water along with fluoride in some Indian populations; however, based on water quality maps, arsenic does not appear to be an issue in the study area.
    - *Direction/magnitude of effect***:** Presence of arsenic would potentially bias away from the null if present with fluoride or toward the null if present in the reference group.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and reliable, that potential co-exposures were not an issue, and that key confounders were addressed.

- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Results were provided for 84 children, but the methods do not indicate how many children were initially selected to participate nor were any exclusion criteria provided. It was noted in the results that 84 children had their groundwater and urine tested, but it was not noted if analyses were restricted to these children or if exposures were assessed in all the children who had IQ measurements. Correspondence with the study author indicated that the main reason for exclusion was a <80% attendance rate, with fluoride and IQ measured on all 84 children who met the criteria.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence of no attrition.

- ***Exposure:***

  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Children in villages were grouped based on fluoride levels that were assessed in groundwater (low F villages versus high F villages). The average concentration of these levels was considered to be the levels in the drinking water with confirmation using urinary fluoride levels. The groundwater samples were selected to cover major parts of the taluka and represent overall groundwater quality. Ten samples were obtained from each village. Fluoride was measured in the groundwater using ion exchange chromatography. Although urine levels were also significantly higher in the high fluoride village, no information was provided on how or when the urinary samples were obtained or how they were measured. However, correspondence with the study author indicated that the groundwater and urine fluoride levels were available for all 84 children indicating that the urine measures were available for the children that had IQ measures. The urine samples were stated to be collected at the same time that the second water sample was collected.
    - *Direction/magnitude of effect***:** Fluoride levels were measured in both the drinking water and urine. Although there is some variability in the measurements, there is no overlap between the two groups and the urine and drinking water levels in the children support each other. Any potential exposure misclassification would be non-differential and direction and magnitude are unknown.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Outcome methods were only noted to be reported in Trivedi *et al.* (2007), which was scored as follows: IQ was measured in the children of both areas using a questionnaire prepared by Professor JH Shah, copyrighted by Akash Manomapan Kendra, Ahmedabad, India, and standardized on the Gujarati population with 97% reliability rate in relation to the Stanford-Binet Intelligence Scale (+ for methods). Blinding or other methods to reduce bias are not reported, but correspondence with the study author indicated that the teachers were blind to the status of fluoride. The teachers administered the tests in the presence of a research fellow. It is not completely clear who scored the tests, but it is assumed the teachers. (+ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcomes were blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***

  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were reasonable (paired sample T-test), but consideration of homogeneity of variance was not reported.
    - *Other potential concerns:* No other threats to internal validity were identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements and the addressing of potential key confounders but was limited by the cross-sectional study design.

Valdez Jimenez *et al.* (2017)

Study Details:

- ***Study design*:** Prospective cohort
- ***Population*:** Infants aged 3–15 months
- ***Study area:*** Durango City and Lagos de Moreno, Jalisco, Mexico
- ***Sample size:*** 65 infants
- ***Data relevant to the review:*** The Bayley Scales of Infant Development II was used to assess Mental Development Index Scale and the Psychomotor Development Index scale in children 3 to 15 month and evaluated for associations with first and second trimester maternal urine fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between maternal urinary fluoride and MDI score during first trimester (adjusted β = −19.05; SE = 8.9) and second trimester (adjusted β = −19.34; SE = 7.46).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were recruited from two endemic areas in Mexico. The study authors do not provide information on the similarities or differences between the two areas nor do they indicate if there were different participation rates. However, recruitment methods were the same. Women receiving prenatal care in health centers located in Durango City and Lagos de Moreno, Jalisco, Mexico were recruited in 2013–2014. Participation rates are not likely to be an issue as characteristics were similar between those who participated and those who did not. Although they did not provide characteristics by area, the characteristics provided do not indicate any differences that may be biased by the selection. Considering the age range for the non-participants, the mean age for non-participants appears to be incorrect (or the age range is incorrect);

however, there does not appear to be a difference that would potentially indicate selection bias.
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited with the same methods in the same time frame, with no evidence of differences or issues with participation/response rates.
- ***Confounding:***
    - *Rating:* Probably high risk of bias (-)
    - *Summary:* Questionnaires were used to obtain information about sociodemographic factors, prenatal history, mother's health status before pregnancy (e.g., use of drugs, vaccines, diseases) and the type of water for drinking and cooking. The marginalization index (MI) was obtained from the National Population Council (CONAPO). Two additional surveys were conducted during the 2nd and 3rd trimester of pregnancy to get information about the mother's health, pregnancy evolution, and sources of water consumption. A survey was also conducted to get information about childbirth (type of birth, week of birth, weight and length of the baby at birth, Apgar and health conditions of the baby during the first month of life). This information was corroborated with the birth certificate. Linear regression models included gestational age, children's age, marginality index, and type of drinking water. Bivariate analysis was conducted on the other factors including child's gender prior to conducting multivariable regression models. Some important confounders were not considered, including parental mental health, IQ, smoking, and potential co-exposures.
    - *Potentially important study-specific confounders:* Arsenic is a potential co-exposure in this area of Mexico.
        - *Direction/magnitude of effect***:** If arsenic were present as a co-exposure it would bias the results away from the null.
    - *Basis for rating:* Probably high risk of bias based on indirect evidence that there is a potential for co-exposure with arsenic that was not addressed.
- ***Attrition:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* Out of the 90 women selected for inclusion in the study, 65 approved the participation of their infants. The authors provide a table of characteristics between women who consented to their children's cognitive evaluation and those that only participated in biological monitoring. There were no significant differences between the groups. There were fewer women who provided urine during the second and third trimesters. All specified children are included in the relevant analyses.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* Fluoride exposure is assessed through morning urine samples and water fluoride levels collected from the children's homes. Sampling methodology is appropriately documented, and water levels were quantified through specific

ion-sensitive electrode assays. QC was described and accuracy was >90%. Urinary fluoride was corrected by specific gravity.

- *Direction/magnitude of effect***:** Not applicable.

- *Basis for rating:* Definitely low risk of bias based on direct evidence that exposure was consistently assessed using well-established methods that directly measured exposure.

- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Neurodevelopment was assessed with the Bayley Scales of Infant Development II (BSDI-II) that was noted to be reliable and valid for evaluating children from 3 months to 5 years of age. The average age of children assessed was 8 months, with a range of 3–15 months) (++ for methods). The study report stated that a trained psychologist who was blinded about the mother's fluoride exposure evaluated the infants at home (++ for blinding). Overall rating for methods and blinding = ++.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported. Table 4 of the study only displays data for trimesters 1 and 2. Although 3rd trimester data were collected, they were not reported, likely because data were only available for 29 subjects. No discussion of this was provided.
  - *Basis for rating:* Probably low risk of bias because, although it appears some data were not reported, it is likely because there were insufficient data and not because the authors were selectively reporting the results.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study.
    - *Other potential concerns:* No other potential concerns were identified. In the peer-review report, NASEM (2020) cited the following as potential concerns: "the large difference in numbers of males and females in the offspring (20 males, 45 females), and apparently incorrect probabilities were reported for age differences between participants and nonparticipants, high rates of cesarean deliveries and premature births among participants (degree of overlap not reported), and incorrect comparisons of observed prematurity rates with national expected rates." However, these concerns were taken in consideration in other domains (***Selection, Confounding)***.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and

outcome blindly assessed, but it is limited by the cross-sectional study design and lack of accounting for potential co-exposures to arsenic.

Wang *et al.* (2012)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 8–13 years (possibly the same population as Xiang *et al.* (2003a))
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Providence, China
- ***Sample size:*** 526 school children
- ***Data relevant to the review:*** Mean IQ and % low IQ (< 80) by total fluoride intake.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower mean IQ in the high fluoride village (92.02 ± 13.00) compared to the control village (100.41 ± 13.21); when high exposure group was broken into 4 exposure groups, a dose-dependent decrease in IQ and increase in % with low IQ observed; significant correlation between total fluoride intake and IQ (r = −0.332); OR for IQ<80 per increase in total fluoride intake=1.106; 95% CI 1.052–1.163).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study appears to be the same population as Xiang *et al.* (2003a) and Xiang *et al.* (2011); however, the study does not cite these studies as providing additional information and numbers of children differ; therefore, it may be a separate analysis on the same villages. The years of testing were not provided so it cannot be determined if study subjects are the same. Two villages, Wamiao and Xinhuai, located 64 km apart in Sihong County, Jiangsu Province were selected for the study. Wamiao is a village in a region with severe endemic fluorosis and Xinhuai is a village in a non-endemic fluorosis region. Neither village has fluoride pollution from coal or industrial sources. Villages were stated to be similar in terms of annual per capita income, transportation, education, medical conditions, the natural environment, and lifestyle. All primary students ages 8–13 years currently in school in either village were surveyed with exclusions noted. Of 243 children from Wamiao, 236 (97.12%) were included, and of 305 children from Xinhuai, 290 (95.08%) were included. No table of subject characteristics was provided.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited using the same methods within the same time frame, with direct evidence that there was no difference in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Logistic regression of low IQ rate and total fluoride intake adjusted for age and sex. Both villages had hand-pumped well water for drinking water, but the authors

do not mention if arsenic was also present in the drinking water. However, a publication by Xiang *et al.* (2013) on this study area indicates that Xinhuai (the low fluoride area) had significantly higher arsenic levels compared to Wamiao (the endemic fluorosis area), which would bias toward the null. Areas were stated to be similar in annual per capita income, transportation, education, medical conditions, the natural environment, and lifestyle; however, no details were provided. This study did not address other co-exposures, but other studies on populations in these villages (Xiang *et al.* 2011, Xiang *et al.* 2003a) indicate that iodine and lead are not concerns.

- *Potentially important study-specific confounders:* Arsenic often occurs in the drinking water along with fluoride in some Chinese populations; however, based on information provided in Xiang *et al.* (2013), arsenic concentrations were higher in the low fluoride area compared to the high fluoride area.
  - *Direction/magnitude of effect*: Presence of arsenic in this study population would potentially bias toward the null.
- *Basis for rating:* Probably low of risk bias because there is indirect evidence that the key confounders are take into account, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency are not attributing to the effect observed in this area.

- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Data are reported for all 526 children noted to be included in the study. There is a slight discrepancy in the reported total number of children from the high-fluoride village and the number of participants from the high-fluoride village between this paper (236 participated of 243 total children) and the 2003 and 2011 publications on the same study population (222 of 238). This discrepancy is not explained but is not expected to appreciably bias the results.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+); Probably high risk of bias (-)
  - *Summary:* **Water fluoride (+ probably low risk of bias):** Exposure was based on drinking water levels and fluoride intake. Residents in the Wamiao village were divided into five groups based on fluoride levels in the drinking water. Clean, dry polyethylene bottles were used to collect 50 mL of drinking water from each student's household and fluoride content was measured.

    Total fluoride intake (- probably high risk of bias): Six families from each of the five Wamiao groups were randomly selected as dietary survey households. Intakes of various foods by each person at each meal and intakes of unboiled water, boiled water, and tea were surveyed for four consecutive days. Methods for food collection were described. Five representative households from each village were selected based on geographic location, population distribution, housing structure, and other conditions. Indoor air samples were collected once daily for five consecutive days; outdoor air was sampled at two points once daily for five days. Methods for determining fluoride

content in samples were noted to follow specific guidelines. Calculation of total fluoride intake was stated to follow Appendix A of the People's Republic of China Health Industry Standard with some details provided. Although it is assumed the method is valid, it was not detailed how each fluoride determination was made for each subject, and it appears that total fluoride intake was determined based on data from select subjects and not all subjects.

- *Direction/magnitude of effect*: There is potential for exposure misclassification based on calculating fluoride intake based on measurements from a few select subjects rather than all subjects. The direction and magnitude of effect cannot be assessed based on the information provided.

- *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure. The intake is probably high risk of bias because there is indirect evidence that the exposure was assessed using methods of unknown validity.

- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* IQ of each child was measured with the Combined Raven's Test for Rural China (CRT-RC) (++ for methods). The test was stated to be administered to the children independently in a school classroom under the supervision of three exam proctors. Testing methods, testing language, and testing conditions were all in strict accordance with the CRT-RC guidebook. Major testing personnel received necessary training by the Psychology Department of East China Normal University. The children undergoing IQ testing and the test scorers were kept double-blinded throughout the testing process. (++ for blinding). Overall rating= ++.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Although it was noted that a logistic regression was used to determine the odds of having low IQ with increasing fluoride intake, no details were provided on any of the other tests conducted. Because this is the same population evaluated in Xiang *et al.* (2003a) and Xiang *et al.* (2011), it is assumed that the same methods were used even if this study population consisted of different children.
    - *Other potential concerns:* None identified.

  - *Basis for rating:* Probably low risk if bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements with blinding at outcome assessment but is limited by the cross-sectional study design and not using individual measurements to calculate fluoride intake. All key confounders were accounted for in the study design or analysis, but there is potential for the presence of arsenic to bias toward the null.

Wang *et al.* (2020a)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** School children aged 7–13 years
- ***Study area:*** Tongxu County, China
- ***Sample size:*** 325 school children
- ***Data relevant to the review:*** Associations between ADHD and other measures of learning disability with urine fluoride concentrations.
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between psychosomatic problems and urinary fluoride (adjusted β = 4.01 [95% CI: 2.74, 5.28]) and increased risk of a T-score > 70 with increasing urinary fluoride (adjusted OR = 1.97 [95% CI: 1.19, 3.27]). No significant associations with ADHD or other measures of learning disability.

Risk of Bias:

- **Author contacts:**
  - Authors were contacted in July of 2020 to obtain additional information for risk-of-bias evaluation. No response was received.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Subjects were recruited in 2017 from Tongxu County, China. Children were selected from four randomly selected primary schools in the area. Selection was based on specified inclusion rules. It was noted that the living habits and diets of the participants from the four schools were well matched, but details were not provided. The area did not have industrial pollution within 1 km of the living environment of the children, and it was noted that the children were not exposed to other neurodevelopmental toxicants (lead, cadmium, arsenic, or mercury). A table of subject characteristics was provided in the study, but not by school or exposure. This is a pilot study, and it is not explicitly stated if all eligible subjects participated in the study. There is no information on participation rates or if they varied by school.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposed groups were recruited using similar methods during the same time frame and that any differences between the exposed groups were considered in the statistical analyses.

- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* It was noted that subjects were well matched in terms of living habits and diets, but there were no specifics provided. It was noted that there was no industrial exposure or exposure to other neurotoxins such as lead, cadmium, arsenic, or mercury. Covariates were collected using a standardized and structured questionnaire completed by the children and their guardians under the direction of investigators, but reliability or validity of the questionnaire was not reported. Information collected included age, gender, weight, height, parental education level, and parental migration (or work as migrant workers). IQ scores evaluated by the Combined Raven's Test-the Rural in China were used to represent basic cognitive function. Models were adjusted for age, BMI, gender, mother and father migration, and urinary creatinine. Adjustments were not made for parental education, race/ethnicity, maternal demographics (e.g., maternal age, BMI), parental behavioral and mental health disorders (e.g., ADHD, depression), smoking (e.g., maternal smoking status, secondhand tobacco smoke exposure), reproductive factors (e.g., parity), iodine deficiency/excess, maternal (and paternal) IQ, quantity and quality of caregiving environment (e.g., HOME score), or SES other than parental migration. There is no evidence to suggest that SES would differ substantially among the four rural schools in the same area of China that were randomly selected.
  - *Potentially important study-specific confounders:* SES.
    - *Direction/magnitude of effect***:** Direction and magnitude is unknown. It was noted that the subjects were matched in terms of living habits and diet and this could be an indication that SES was not different among the groups, but details were not provided.
  - *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key confounders are considered, that the methods for collecting the information were valid and reliable, and that co-exposure to arsenic is not an issue in this area.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Data are complete. It was noted that there were 325 subjects included and results were available on all subjects.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that there was no attrition.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Spot urine samples were collected from each child in the early morning into cleaned polyethylene tubes. Fluoride concentrations were measured using fluoride ion-selective electrode (with reference to Ma *et al.* (2017); however, that reference cites Zhou *et al.* (2012). Therefore, no QC methods or LODs were available. Fluoride concentrations were creatinine-adjusted.
    - *Direction/magnitude of effect***:** Spot urine samples only account for recent exposure. Although this could cause there to be some exposure misclassification, the number of subjects should help dilute any issues with the non-differential misclassification.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using acceptable methods that provide individual levels of exposure.
- ***Outcome:***
  - *Rating:* Probably high risk of bias (NR)
  - *Summary:* Children's behavior was assessed by the Chinese version of the Conners' Parent Rating Scale-Revised (CPRS-48). The homogeneity reliability of Cronbach α in the Chinese version of CPRS-48 was 0.932; the correlation of Spearman-brown split-half was 0.900; and the retest reliability of total score was 0.594. Raw scores for each subscale are converted into sex- and age-adjusted T-scores within a mean ± standard deviation (SD) of 50 ± 10. The guardians independently completed the CPRS-48 according to the instruction manual under the direction of trained investigators (++ for methods). Blinding is not reported. Although it is unlikely that the outcome assessors were aware of the fluoride levels in the urine, it is unclear if subjects were selected based on areas with endemic fluoride or if parents were aware of fluoride concentrations in the areas. (NR for blinding). Overall rating for methods and blinding = NR.
  - *Basis for rating:* Probably high risk of bias based on no information provided to indicate that the outcome was blindly assessed.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes in the abstract, introduction, and methods are reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Multiple linear regression models were used to assess the fluoride association with each behavioral outcome. Logistic regression was used to assess the risk of behavioral problems due to fluoride exposure, but what they used to delineate a behavioral problem was not specified. Sensitivity analyses were performed.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and no other potential threats to risk of bias were identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include individual exposure measurements, but it is limited by the cross-sectional study design and lack of details on blinding of the outcome assessment. All key confounders were considered in the study design or analysis.

Wang *et al.* (2020b)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** School children aged 7–13 years
- ***Study area:*** Tianjin City, China (possibly a subset of the children from Yu *et al.* (2018))
- ***Sample size:*** 571 school children
- ***Data relevant to the review:*** IQ scores by urine and water fluoride levels.
- ***Reported association with fluoride exposure:*** Yes: Significant correlations between IQ score and water fluoride (adjusted β = −1.587 per 1-mg/L increase) and urinary fluoride (adjusted β = −1.214 per 1-mg/L increase).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Subjects were from a cross-sectional study conducted in 2015, but no citation was provided on this cohort (presumably the Yu *et al.* (2018) cohort). It was noted that the subjects in that cohort were from districts with historically high or normal fluoride levels. Subjects for this study were selected by using a stratified and multistage random sampling approach. Brief description was provided. The study area consisted of three historically high fluoride areas and four nonendemic areas. A flow diagram was provided for inclusion and exclusion, but this detail was given for all children and not by area. Therefore, it cannot be determined if the participation differed by area. However, there was a 93% recruitment rate, and the 13 excluded due to missing data are not likely excluded due to exposure. Detailed characteristics of the study population are provided. Exclusion criteria included: "children who had congenital or acquired diseases affecting intelligence, or a history of cerebral trauma and neurological disorders, or those with a positive screening test history (like hepatitis B virus infection, Treponema palladium infection and Down's syndrome) and adverse exposures (smoking and drinking) during maternal pregnancy, prior diagnosis of thyroid disease, and children who had had missing values of significant factors (2.2%) were also excluded."
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposed groups were recruited using similar methods during the same time frame and that any differences between the exposed groups were accounted for in the statistical analyses.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Study authors noted that the study areas are not exposed to other neurotoxins such as lead, arsenic, or mercury nor were they iodine-deficient. Final models included child's age, child's gender, child's BMI, maternal and paternal education, household income, and low birth weight. Other potential confounders that were considered is unclear as they only noted that the confounders were selected based on current literature. Reasons for exclusion included history of disease affecting intelligence, history of trauma or neurological disorders, positive screening test history,

or exposures such as smoking or drinking during pregnancy. Information was obtained by questionnaire or measurements. Variables such as parental BMI, behavioral and mental health disorders, IQ, and quantity and quality of the caregiving environment were not addressed.

- *Potentially important study-specific confounders:* All key confounders were considered in this study.
  - *Direction/magnitude of effect*: Not applicable.
- *Basis for rating:* Probably low risk of bias because there is direct evidence that the key confounders are taken into account, indirect evidence that the methods for collecting the information were valid and reliable, and co-exposure to arsenic is not an issue in this area.

- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* A detailed chart of the recruitment process is presented. The study had a 93% recruitment rate and only 2.2% of subjects with missing data for certain covariates were excluded.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Children provided spot urine samples, presumably at the time of examination. Water samples were randomly collected from public water supplies in each village. Fluoride concentrations were analyzed using fluoride ion-selective electrode according to the national standardized method in China. There is no indication if the urine samples accounted for dilution.
    - *Direction/magnitude of effect*: Not accounting for dilution could cause there to be some exposure misclassification. The direction and magnitude would depend on where the differences occurred.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using acceptable methods that provide individual levels of exposure.
- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Assessments of IQ scores were conducted by graduate students at the School of Public Health, Tongji Medical College at the Huazhong University of Science and Technology. Each team member was assigned a single task, meaning that only one person would have conducted the IQ tests. A Combined Raven's Test for Rural China was used. Therefore, the test was appropriate for the study population (++ for method). It was note that the examiner was trained and blind to the exposure (++ for blinding). Overall = ++
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.

- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes in the abstract, introduction, and methods are reported in sufficient details.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses were appropriate and no other threats to internal validity were identified. Logistic and multivariate regression models accounting for potential confounders were used. Results are presented as betas or odds ratios and 95% confidence intervals. Regression diagnostics were conducted for all models, including examination of multicollinearity, heteroscedasticity, and influential observations. Mediation and interaction analyses were appropriate. The stratified and multistage random sampling approach for subject selection and the fact that selected villages are similar in population and general demographic characteristics helps to ensure that there was no need to account for village-level effects even when the analysis used water samples from the village.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the statistical analyses were appropriate and no other potential threats to risk of bias were identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements but is limited by the cross-sectional study design and lack of accounting for urine dilution. All key confounders were considered in the study design or analysis.

Xiang *et al.* (2003a)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 8–13 years
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Providence, China
- ***Sample size:*** 512 school children
- ***Data relevant to the review:*** Comparison of IQ (mean and distribution) between Wamiao County (a severe endemic fluorosis area) and Xinhuai County (non-endemic fluorosis area); additional breakdown of the Wamiao area into 5 water fluoride exposure groups.
- ***Reported association with fluoride exposure:*** Yes: Significant dose-related effect of drinking water fluoride on IQ score based on quintile levels with significantly lower IQ scores observed with water fluoride levels of 1.53 mg/L or higher. Pearson correlation coefficient of −0.164 with urinary fluoride. IQ scores for children in the non-endemic region (100.41 ± 13.21) were

significantly higher than the endemic region (92.02 ± 13.00). The lower-bound confidence limit benchmark concentration (BMCL) of 1.85 mg/L was calculated.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Two villages, Wamiao and Xinhuai, located 64 km apart in Sihong County, Jiangsu Province were selected for this study, which was conducted between September and December 2002. Wamiao is located in a severe fluorosis endemic area, and Xinhuai is located in a non-endemic fluorosis area. Neither village has fluoride pollution from burning coal or other industrial sources. All eligible children in each village were included; children who had been absent from either village for 2 years or longer or who had a history of brain disease or head injury were excluded. In Wamiao, 93% of the children (222 out of 238) were included for the study, while in Xinhuai, 95% were included (290 out of 305). The children in Wamiao were divided into five subgroups according to the level of fluoride in their drinking water: <1.0 mg/L (group A), 1.0–1.9 mg/L (group B), 2.0–2.9 mg/L (group C), 3.0–3.9 mg/L (group D), and >3.9 mg/L (group E). Children in Xinhuai (0.18–0.76 mg F/L in the drinking water) served as a control group (group F). Demographic characteristics are not presented, and statistical analyses are not adjusted, but mean IQ scores are stratified by child's age, child's gender, family income, and parental education.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited using the same methods within the same time frame, with direct evidence that there was no difference in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Although information was stated to be collected on personal characteristics, medical history, education levels of the children and parents, family SES, and lifestyle, only child's gender, child's age, family income, and parental education were addressed. Other potential co-exposures, such as arsenic, were not addressed. A separate publication in 2003 [(Xiang *et al.* 2003b), letter to the editor], indicated that blood lead levels were not significantly different between the two areas. Although arsenic was not addressed specifically in this publication, Xiang *et al.* (2013) measured both fluoride and arsenic in the Wamiao and Xinhuai areas. Xinhuai (the low fluoride area) had significantly higher arsenic levels compared to Wamiao (the endemic fluorosis area). This is likely to bias toward the null; however, the study observed a significantly lower IQ score in the endemic fluorosis area. Iodine was tested in a subset of the children and found not to be significantly different between the two groups.
  - *Potentially important study-specific confounders:* Arsenic often occurs in the drinking water along with fluoride in some Chinese populations; however, based on information provided in Xiang *et al.* (2013), arsenic concentrations were higher in the low fluoride area compared to the high fluoride area.

        - *Direction/magnitude of effect*: Presence of arsenic in this study population would potentially bias towards the null.
    - *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key confounders are taken into account, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency are not attributing to the effect observed in this area.
- ***Attrition:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* Data are complete. IQ results were reported for all 512 children included in the study (222 in the endemic area and 290 in the nonendemic area).
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that there was no attrition.
- ***Exposure:***
    - *Rating:* Probably low risk of bias (+)
    - *Summary:* Exposure was based on drinking water and urinary levels of fluoride. The two study areas were selected to reflect a severe endemic area and a nonendemic area. Drinking water was collected from wells and early-morning spot urine samples were collected from a randomly-selected subsample of children. Both water and urine samples were measured using fluoride ion-selective electrode, but no quality control was discussed. Both absolute and creatinine-adjusted urine results were reported.
        - *Direction/magnitude of effect*: There is potential for exposure misclassification because only current levels were assessed. Migration of subjects in or out of the area was not assessed, but the study authors noted that, if the children had been absent from the village for 2 or more years, they were excluded. Misclassification would likely be non-differential, which could bias the results in either direction.
    - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* The IQ of each child was measured with the Combined Raven's Test for Rural China (CRT-RC) (++ for methods). The test was stated to be administered to the children independently in a school classroom, in a double-blind manner, under the supervision of an examiner and two assistants, and in accordance with the directions of the CRT-RC manual regarding test administration conditions, instructions to be given, and test environment. (++ for blinding). Overall rating= ++
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.

  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Data were stated to be analyzed using SAS without reporting the tests conducted. Results provided in the tables indicate that a t-test was conducted, but it was not reported that homogeneity of variance was tested or confirmed. In addition, correlations were tested with Pearson's correlation.
    - A potential concern raised by the NASEM (2020) peer review was the lack of accounting for relationships in exposure between persons from the same village. However, only two villages were included, and the analyses consisted of village-level comparisons; hence, accounting for clustering was not possible. Without controlling for village effects and given the large differences in fluoride concentrations and IQ between villages, the apparent dose-response relationship could be due to a village effect in addition to a fluoride effect. However, the dose-response relationship is still present within the "exposed" village, diminishing the concern for a village-only effect.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that statistical analyses were appropriate and that there were no other threats to risk of bias.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements and outcomes blindly assessed but is limited by the cross-sectional study design and lack of accounting for urine dilution. All key confounders were considered in the study design or analysis, but there is potential for the presence of arsenic to bias toward the null.

Xiang *et al.* (2011)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 8–13 years (same population as Xiang *et al.* (2003a))
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Providence, China
- ***Sample size:*** 512 school children
- ***Data relevant to the review:*** Mean IQ scores and odds ratio for having an IQ < 80 presented by serum fluoride quartiles.
- ***Reported association with fluoride exposure:*** Yes: Significant trend on association between quartiles of serum fluoride and children's IQ score < 80 (adjusted ORs for Q1 and Q2; Q1 and Q3; and Q1 and Q4, respectively: 1; 2.22 [95% CI: 1.42, 3.47]; and 2.48 [95% CI: 1.85, 3.32]); significant effects at ≥ 0.05 ppm fluoride.

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study population is the same as that was used in the Xiang *et al.* (2003a) study, but a few more measurements were available and different analyses were conducted. The comparison population is considered the same as previously based on the study populations being recruited from similar populations, using similar methods, during the same time frame. Demographic characteristics were not provided.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the exposure groups were similar and were recruited using the same methods within the same time frame, with direct evidence that there was no difference in participation/response rates.
- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* As was noted in the 2003 publication, information was collected on personal characteristics, medical history, education levels in the children and parents, family SES, and lifestyle. In the logistic regression model, age and gender were adjusted in the analysis. In the previous report, no significant associations were observed between groups for family income and parents' education. Urinary iodine and blood lead levels were also stated to be measured and were noted not to be significantly different between the groups. Although the iodine levels were reported in the previous publication, the lead levels were not reported nor were the methods. Lead information is reported in a letter to the editor (Xiang *et al.* (2003b)) and was not significantly different between the areas. Although arsenic was not addressed specifically in this publication, Xiang *et al.* (2013) measured both fluoride and arsenic in the Wamiao and Xinhuai areas. Xinhuai (the low fluoride area) had significantly higher arsenic levels compared to Wamiao (the endemic fluorosis area). This is likely to bias toward the null; however, the study observed a significantly lower IQ score in the endemic fluorosis area and with increasing serum fluoride.
  - *Potentially important study-specific confounders:* Arsenic often occurs in the drinking water along with fluoride in some Chinese populations; however, based on information provided in Xiang *et al.* (2013), arsenic concentrations were higher in the low fluoride area compared to the high fluoride area.
    - *Direction/magnitude of effect***:** Presence of arsenic in this study population would potentially bias toward the null.
  - *Basis for rating:* Probably low of risk bias because there is indirect evidence that the key confounders are taken into account, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency are not attributing to the effects observed in this area.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Data are reported for all 512 children noted to be included in the study.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that there was no attrition.

- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Fluoride levels were measured in serum with a fluoride ion-selective electrode. A fasting venous blood sample was used. No details are provided on validation (including correlation with drinking water levels) or QA. Children who did not reside in their village for at least 2 years were excluded. Results were provided in quartiles, but they combined the lower two quartiles with results ranging from <0.05 mg/L to >0.08 mg/L.
    - *Direction/magnitude of effect***:** Serum fluoride may not be the best estimate for exposure. There is potential for exposure misclassification because only current levels were assessed. Migration of subjects in or out of the area was not assessed, but the study authors noted that, if the children had been absent from the village for 2 or more years, they were excluded. Misclassification would likely be non-differential, which could bias results in either direction.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* IQ was assessed as part of the 2003 evaluation. IQ was measured with the Combined Raven's Test for Rural China which is appropriate for this population (++ for methods). Although this study does not provide details, the original study article from 2003 provides specific details. The study authors indicate in the 2003 publication that the tests were conducted in a double-blind manner and these are the same results and population (++ for methods). Overall rating=++
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses conducted were appropriate for the study. Chi square tests were used to compare categorical variables, and logistic regression was used to evaluate the association between serum fluoride levels and risk of low IQ. A potential concern raised by the NASEM (2020) peer review was the lack of accounting for relationships in exposure between persons from the same village. However, only two villages were included, and the analyses consisted of village-level comparisons; hence, accounting for clustering was not possible. Without controlling for village effects and given the large differences in

fluoride concentrations and IQ between villages, the apparent dose-response relationship could be due to a village effect in addition than a fluoride effect. However, the dose-response relationship is still present within the "exposed" village, diminishing the concern for a village-only effect.

- *Other potential concerns:* None identified.

- *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.

- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements with blinding at outcome assessment but is limited by the cross-sectional study design and use of serum concentrations. All key confounders were considered in the study design or analysis, but there is potential for the presence of arsenic to bias toward the null.

Yu *et al.* (2018)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 7–13 years
- ***Study area:*** Tianjin City, China
- ***Sample size:*** 2,886 school children
- ***Data relevant to the review:*** IQ for normal (≤ 1 mg/L) versus high (> 1 mg/L) water fluoride; betas for IQ score by water and urine fluoride groupings; ORs by IQ category using water and urine fluoride levels.
- ***Reported association with fluoride exposure:*** Yes: Significant difference (p = 0.036) in mean IQ scores in high (106.4 ± 12.3) versus normal (107.4 ± 13.0) water fluoride areas. Distribution of IQ scores was also significantly different (p = 0.003); every 0.5-mg/L increase in water fluoride (between 3.40 and 3.90 mg/L) was associated with an IQ score 4.29 points lower (95% CI: −8.09, −0.48).

Risk of Bias:

- **Author contacts:**
  - Authors were contacted in September 2018 to obtain additional information for the risk-of-bias evaluation.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* School children (2,886), aged 7–13 years, were recruited from the rural areas of Tianjin City, China. After exclusion, 1,636 children were assigned to the "normal-fluoride" exposure group and 1,250 were assigned to the "high-fluoride" exposure group based on a cut-off water fluoride level of 1.0 mg/L. A multi-stage random sampling technique, stratified by area, was performed to select representative samples among local children who were permanent residents since birth. Detailed characteristics of the study population are provided. Exclusion criteria included: 1) children who had

congenital or acquired diseases affecting intelligence, 2) children with a history of cerebral trauma and neurological disorders, 3) children with a positive screening test history (like hepatitis B virus infection, Treponema palladium infection and Down's syndrome), and 4) children with adverse exposures (smoking and drinking) during maternal pregnancy. A table of characteristics was provided by fluoride level with differences adjusted in the analysis.

- *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposed groups were recruited using similar methods during the same time frame and that any differences between the exposed groups were considered in the statistical analyses.

- ***Confounding:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Demographic data were collected by trained investigators during a face-to-face interview with the recruited children and their parents. Questionnaires were not stated to be validated. The developmental status of the children was further assessed by calculation of BMI, and all measurements were conducted by nurses based on recommended standard methods. Variables that presented differential distribution between the normal-fluoride and high-fluoride exposure groups were adjusted in the linear regression analysis of IQ data and included age, sex, paternal and maternal education levels, and low birth weight. Children exposed to smoking in utero were excluded from the study. Sensitivity analyses were conducted by modifying covariates adjusted in multivariable models among demographics (age and sex); development (BMI); socioeconomics (maternal education, paternal education, and household income); history of maternal disease during pregnancy (gestational diabetes, malnutrition, and anemia); and delivery conditions (hypoxia, dystocia, premature birth, post-term birth, and low birth weight). None of the study sites selected were in areas endemic for iodine deficiency disorders nor were other potential neurotoxins like lead, arsenic, and mercury present. Variables such as parental BMI and behavioral and mental health disorders were not addressed.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect***:** Not applicable.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that methods of obtaining the information were valid and reliable and direct evidence that all key confounders and co-exposures were addressed.
- ***Attrition:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* There were 1,636 children assigned to the "normal-fluoride" exposure group based on water fluoride, and 1,250 children were assigned to the "high-fluoride" exposure group. Exclusion from the original group of 2,886 children was adequately described. A total of 2,380 children provided urine samples. There is no indication that the data presented excludes any additional children or urine samples, but results do not indicate a sample size for all results.

  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exclusion of subjects from analyses was adequately addressed, and reasons were documented when subjects were removed from the study or excluded from analyses.
- ***Exposure:***
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* According to the annual surveillance data from the CDC, the drinking water sources and water fluoride concentrations in each village had remained at stable levels over the past decade. During the investigation, water samples were collected randomly from the public water supplies in each village. Spot (early-morning) urine samples from every child and water samples from each village were collected in pre-cleaned, labeled polythene tubes and transported to the lab within 24 hours while frozen. Samples were stored at −80°C until analysis. Concentrations of fluoride ions (mg/L) were analyzed using the national standardized ion-selective electrode method in China; the detection limit was 0.01 mg/L. Samples were diluted with an equal volume of total ionic strength adjusted buffer (TISAB) of pH 5–5.5 for optimal analysis. Double-distilled deionized water was used throughout the experiment. There is no reporting of any QC methods.
    - *Direction/magnitude of effect*: Spot urine samples may lead to non-differential exposure misclassification. The large population size likely dilutes any potential effects of occasional misclassification. Because the drinking water sources of fluoride had been noted to be stable for the past decade and the children were 13 years or younger, there would only be exposure misclassification if there was a lot of migration between areas.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* IQ scores were measured using the second edition of Combined Raven's Test-The Rural in China (CRT-RC2) for children aged 7–13 years (++ for methods). The test was completed by each participant within 40 minutes according to the instruction manual. For each test, 40 children were randomly allocated to one classroom to take the test independently under the supervision of four trained professionals. There is no mention of whether the evaluators were blinded to the fluoride group of each child (normal vs. high fluoride) or whether there were steps taken to ensure consistency in scoring across the evaluators. It is also not clear if the 40 children randomly assigned to the classroom were specific to the village or if a local center was used. Correspondence with the study authors indicated that the four professionals worked together throughout the examination without knowledge of the child's fluoride exposure (++ for blinding).
  - *Basis for rating:* Definitely low risk of bias based on the direct evidence that the outcome was blindly assessed using instruments that were valid and reliable.
- ***Selective Reporting:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods are reported in sufficient detail.

  - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements with blinding at outcome assessment but is limited by the cross-sectional study design and lack of accounting for urine dilution. All key confounders including potential co-exposures were considered in the study design or analysis.

Zhang *et al.* (2015b)

Study Details:

- ***Study design*:** Cross-sectional
- ***Population*:** Children aged 10–12 years
- ***Study area:*** Tianjin City, China
- ***Sample size:*** 180 children
- ***Data relevant to the review:*** IQ by control and high fluoride groups; IQ correlations with water, serum, or urinary fluoride levels; betas for IQ with urinary fluoride levels (by genotypes)
- ***Reported association with fluoride exposure:*** Yes: Significant correlation between IQ score and serum fluoride (r = –0.47) and urinary fluoride (r = –0.45); significant difference in IQ score for high-fluoride area (>1 mg/L; 102.33 ± 13.46) compared with control area (<1 mg/L; 109.42 ± 13.30).

Risk of Bias:

- **Author contacts:**
  - Authors were not contacted for additional information because it was not necessary.
- ***Population selection:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary*: Subjects were similar and recruited during the same time frame using the same methods. Authors recruited schoolchildren from a high fluoride area (1.40 mg/L) and a control area (0.63 mg/L) in Tianjin City, China. In accordance with the principles of matching social and natural factors such as educational standard, economic situation, geological environments as much as possible, two areas with different fluoride concentrations in the groundwater were selected by a stratified cluster random sampling of this region. A total of 180 5th grade children aged 10 to 12 years from two primary schools located 18 km apart in the Jinnan District were recruited—Gegu Second

Primary School (from an endemic fluorosis area) and Shuanggang Experimental Primary School (from a non-endemic fluorosis area). The areas are not affected by other drinking water contaminants, such as arsenic or iodine. All subjects were unrelated ethnic Han Chinese and residents in Tianjin with similar physical and mental health status. The authors excluded subjects with known neurological conditions including pervasive developmental disorders and epilepsy. Descriptive statistics of the study population are presented by exposure group in Table 1 of the study. A number of potential differences are taken into account in the statistical analyses.

- *Basis for rating*: Definitely low risk of bias based on direct evidence that the exposure groups were similar and recruited using similar methods during the same time frame.

- **Confounding:**
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Covariates included in the statistical models were child's age, child's gender, educational levels of parents, drinking water fluoride (mg/L), and levels of thyroid hormones (T3, T4, and TSH). Authors report that the study areas are not affected by other contaminants such as arsenic or iodine and residents were of similar physical and mental health status. Other important confounders (maternal demographics, smoking, reproductive health) were not considered. Covariate data were obtained from a study questionnaire.
  - *Potentially important study-specific confounders:* All key confounders were considered in this study.
    - *Direction/magnitude of effect*: Not applicable.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the methods used to collect the information were valid and reliable and direct evidence that key confounders including potential co-exposures were addressed.
- ***Attrition:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Results are complete for the 180 children selected for the study.
  - *Basis for rating:* Definitely low risk of bias based on direct evidence that there was no attrition.
- ***Exposure:***
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Drinking water samples (10 mL) were collected from the tube wells of each child's household. Three fasting venous blood samples were also collected. Urine samples were collected in the early morning before breakfast. Fluoride contents in drinking water (W-F), serum (S-F), and urine (U-F) were measured using an ion analyzer EA940 with a fluoride ion-selective electrode (Shanghai constant magnetic electronic technology Co, Ltd, China) according to the China standard GB 7484-87. All reference solutions for the fluoride determinations were double-deionized water. Parallel samples were set for determination and averages were taken. The quantitation limits of this method for W-F, S-F, and U-F were 0.2, 0.012, and 0.5 mg/L, respectively. Recovery rates for this method were in the range of 94.3%–106.4%. The intra- and inter-assay coefficients of variation for fluoride were 2.7% and 6.7%, respectively. Dilution of the urinary fluoride was not addressed.

        - *Direction/magnitude of effect*: Not applicable.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that the exposure was consistently assessed using well-established methods that directly measured exposure.
- ***Outcome:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* A Combined Raven's Test for Rural China (CRT-RC) was taken to evaluate the IQ of each child (++ for methods). The study report stated that all tests were administered at school by a trained examiner who was masked to participants' drinking water fluoride levels (++ for blinding). Overall rating for methods and blinding=++.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was blindly assessed using instruments that were valid and reliable in the study population.
- ***Selective Reporting:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:* All results outlined in the abstract, introduction, and methods sections were reported in sufficient detail.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that all measured outcomes were reported.
- ***Other potential threats:***
    - *Rating:* Definitely low risk of bias (++)
    - *Summary:*
        - *Statistical analyses:* Statistical methods are very well-documented including testing for normality of the data.
        - *Other potential concerns:* None identified.
    - *Basis for rating:* Definitely low risk if bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats to risk of bias identified.
- ***Basis for classification as lower risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, outcomes blindly assessed, and assessment of potential key confounders including potential co-exposures.

Appendix 5. Results of Fluoride Meta-analyses

What is the strength of the relationship between exposure to fluoride and children's IQ?

Aim 1. To update existing meta-analyses with additional studies

Approach

The approach used to perform a meta-analysis of the associations between exposure to fluoride and children's IQ levels was outlined in the associated protocol (https://ntp.niehs.nih.gov/go/785076). Details are presented below.

The mean-effect meta-analysis included studies that reported effect estimates as mean outcome measures and included measures of uncertainty such as standard deviation (SD), standard error (SE), 95% CI, and number of subjects (N) for at least one exposed and one reference exposure group. If results from multiple exposure groups were reported within a single study, the highest exposure group was considered the "exposed" group and the lowest exposure group was considered the "reference" group. A sensitivity analysis was performed to evaluate the impact of using any exposed group compared to the reference group (**Figure A-25**). This was accomplished by combining the information from the exposure groups using the approach outlined in the Cochrane Handbook for Systematic Reviews (Higgins *et al.* 2019).

When results were not reported for gender-specific groups or age-specific subgroups (<10, ≥10), they were calculated (if possible) by combining groups, following the approach outlined in the Cochrane Handbook for Systematic Reviews (Higgins *et al.* 2019). Similarly, when only mean effects, Ns, and p-values for differences between groups are reported (Lin *et al.* 1991), SDs were calculated using the SE and t-statistic (assuming equal variances) (Higgins *et al.* 2019).

The meta-analysis pooled the standardized mean difference and corresponding 95% CI using a random-effects model. Heterogeneity was assessed by Cochran's Q test (Cochran 1954) and the I^2 statistic. Forest plots were used to display results and to examine possible heterogeneity between studies. Potential publication bias was assessed by developing funnel plots and performing Egger regression on the estimates of effect size (Begg and Mazumdar 1994, Egger *et al.* 2008, Egger *et al.* 1997). If publication bias was believed to be present, trim-and-fill methods (Duval and Tweedie 2000a, b) were used to estimate the number of missing studies affected by publication bias, assess the effect of those studies on the effect estimate, and predict the impact of the hypothetical "missing" studies. To investigate sources of heterogeneity, subgroup analyses were performed by risk-of-bias evaluation, gender, age group, country, type of intelligence assessment, and type of exposure.

There were 46 studies included in the mean-effect meta-analysis (see **Table A5-1**). **Table A-2** presents information on the studies excluded from the analysis because of missing information on the number of subjects and/or the mean or variance of the outcome. Other studies were excluded because of overlapping study populations. For studies with overlapping populations (i.e., multiple studies that use the same cohort), results were selected with the most information considering the following factors: exposure metric, exposure range, exposure period, number of subjects, and statistical adjustment for potential confounders. Other studies were excluded from the mean-effect meta-analysis because they reported individual-level effects.

Summary Results

Table A5-1. Pooled SMDs and 95% CIs for Children's IQ Score and Exposures to Fluoride

Analysis	Number of Studies	SMD (95% CI)	Heterogeneity	
			p-value	I^2
Overall Effect	46	−0.50 (−0.61, −0.39)	<0.001	89%
Subgroup Analyses				
Risk of Bias				
Lower	9	−0.31 (−0.52, −0.10)	<0.001	87%
Higher	37	−0.56 (−0.68, −0.43)	<0.001	88%
Gender[1]				
Males	12	−0.78 (−0.99, −0.56)	<0.001	75%
Females	11	−0.65 (−0.85, −0.45)	0.001	66%
Age Group				
<10 years[1]	10	−0.55 (−0.79, −0.30)	<0.001	83%
≥ 10 years	11	−0.58 (−0.78, −0.38)	<0.001	76%
Country				
China	31	−0.44 (−0.54, −0.33)	<0.001	86%
India	9	−1.02 (−1.54, −0.51)	<0.001	93%
Iran	4	−0.68 (−0.99, −0.38)	0.077	56%
Assessment Type				
CRT-RC tests	23	−0.36 (−0.47, −0.26)	<0.001	81%
Non-CRT-RC tests	23	−0.67 (−0.87, −0.47)	<0.001	90%
Raven's tests	10	-0.76 (−1.10, −0.43)	<0.001	91%
Other tests	13	−0.62 (−0.88, −0.35)	<0.001	90%
Exposure Type				
Water fluoride	28	−0.45 (−0.57, −0.34)	<0.001	85%
Dental fluorosis	7	−0.99 (−1.57, −0.41)	<0.001	96%
Other exposures[2]	11	−0.47 (−0.67, −0.27)	<0.001	85%
Sensitivity Analysis				
Any exposure *vs.* reference	46	−0.46 (−0.58, −0.35)	<0.001	92%
Previous Meta-analyses				
Duan *et al.* (2018)	26	−0.52 (−0.62, −0.42)	<0.001	69%
Choi *et al.* (2012)	27	−0.45 (−0.56, −0.34)	<0.001	80%

Notes:
CI = confidence interval; CRC-RC = Combined Raven's Test–The Rural edition in China; SMD = standardized weighted mean difference

[1]An *et al.* (1992) includes 10-year-old children in the <10 age group (7−10 years reported).

[2]Includes iodine (Ren *et al.* 1989 [translated in Ren *et al.* 2008], Lin *et al.* 1991, Wang *et al.* 2001); arsenic (Zhang *et al.* 1998, Wang *et al.* 2007); aluminum (Sun *et al.* 1991); and non-drinking water fluoride (i.e., fluoride from coal burning (Guo *et al.* 1991 [translated in Guo et al. 2008a], Li *et al.* 1994 [translated in Li et al. 2008b], Li *et al.* 1995, Wang *et al.* 1996 [translated in Wang et al. 2008b], Wang *et al.* 2005, Li *et al.* 2009, Bai *et al.* 2014)).

Overall Effect (Main Analysis)

For the group-level exposure meta-analysis, a comparison on the mean outcome measure (IQ score) was conducted across two exposure groups ("exposed" and "reference"). The random-effects pooled SMD estimated from the 46 studies included in the meta-analysis was −0.50 (95% CI: −0.61, −0.39). There was evidence of heterogeneity (I^2 = 89%, p < 0.001; **Table A5-1** and **Figure A5-1**) and publication bias (funnel plot and Egger's test p < 0.001, Begg's test p = 0.08; **Figure A5-2** and **Figure A5-3**). Eliminating publication bias through trim-and-fill analysis continued to support the finding that exposure to fluoride is associated with lower IQ in children, with an adjusted pooled effect estimate of −0.42 (95% CI: −0.54, −0.30) (**Figure A5-4** and **Figure A5-5**).

Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017) studies were excluded from the main analysis due to uncertainties about the way IQ assessments for children were performed in those studies. A sensitivity analysis was conducted that included these studies (**Figure A-35**).

Subgroup Analyses

Risk of Bias

Subgroup analysis by risk-of-bias evaluation showed that exposure to fluoride is associated with lower IQ scores in children for both higher and lower risk-of-bias studies (**Figure A5-6**), with a more severe effect for the higher risk-of-bias studies. The funnel plots and Egger's and Begg's tests of publication bias showed evidence of publication bias only among higher risk-of-bias studies (**Figure A-1**, **Figure A-2**). Eliminating publication bias through trim-and-fill analysis continued to support that exposure to fluoride is associated with lower IQ scores in children, with an adjusted pooled effect estimate of −0.35 (95% CI: −0.50, −0.21) (**Figure A-4**). There was no evidence of publication bias among lower risk-of-bias studies.

Gender

Subgroup analysis by gender showed that exposure to fluoride is associated with lower IQ scores in both males and females (**Figure A5-11**). There was a slight suggestion of publication bias in the Egger's test for males, but not females (**Figure A-5** and **Figure A-6**). Eliminating publication bias through trim-and-fill analysis continued to support that exposure to fluoride is associated with lower IQ scores in males, with an adjusted pooled SMD estimate of −0.68 (95% CI: −0.90, −0.46) (**Figure A-8**).

Age Group

Subgroup analysis by age group showed that exposure to fluoride is associated with lower IQ scores in children regardless of age group (<10 years or ≥10 years) (**Figure A5-12**). The funnel plots and Egger's and Begg's tests of publication bias showed evidence of publication bias in children younger than 10 years old (**Figure A-9** and **Figure A-10**). Eliminating publication bias through trim-and-fill analysis for studies in children younger than 10 years old continued to support that exposure to fluoride is associated with lower IQ in children, with an adjusted pooled effect estimate of −0.55 (95% CI: −0.79, −0.30) (**Figure A-10**). There was no suggestion of publication bias in the subgroup analyses for children 10 years old and older.

Country

Subgroup analysis by country showed that exposure to fluoride is associated with lower IQ scores in children in China, India, and Iran (**Figure A5-13**), with the largest effect in India. A funnel plot with the SEs of the SMD plotted against the SMD from each study showed slight evidence of publication bias in India (**Figure A-11**). In addition, Egger's and Begg's tests of publication bias revealed evidence of publication bias for studies in India (p < 0.001, **Figure A-12**). Eliminating publication bias through trim-and-fill analysis for studies in India continued to support that exposure to fluoride is associated with

lower IQ in children, with an adjusted pooled effect estimate of −1.49 (95% CI: −2.15, −0.83) (**Figure A-13**). There was no suggestion of publication bias in the subgroup analyses for China or Iran.

Assessment Type

Subgroup analysis by assessment type showed that exposure to fluoride is associated with lower IQ scores in children tested using non-CRT-RC tests than using CRT-RC tests (**Figure A5-14**). The funnel plots and Egger's and Begg's tests of publication bias showed evidence of publication bias only among non-CRT-RC and Raven's tests (**Figure A-15** and **Figure A-16**). Eliminating publication bias through trim-and-fill analysis for studies with Raven's tests continued to support that exposure to fluoride is associated with lower IQ scores in children, with an adjusted pooled SMD estimate of −1.22 (95% CI: −1.68, −0.75) (**Figure A-19** and **Figure A-20**). There was no suggestion of publication bias in the subgroup analysis for studies using CRT-RC or other types of tests.

Exposure Type

Subgroup analysis by exposure type showed that exposure to fluoride is associated with lower IQ scores in children in studies that reported mean effects by fluoride exposure type (**Figure A5-15**). The funnel plots and Egger's test of publication bias showed evidence of publication bias for water fluoride and dental fluorosis (**Figure A-21** and **Figure A-22**). Eliminating publication bias through trim-and-fill analysis continued to support that exposure to fluoride in water is associated with lower IQ scores in children, with an adjusted pooled SMD estimate of −0.42 (95% CI: −0.53, −0.30) (**Figure A-23** and **Figure A-24**). There was no suggestion of publication bias in the subgroup analysis for studies with other exposures or non-drinking water fluoride exposures.

Overall Analysis

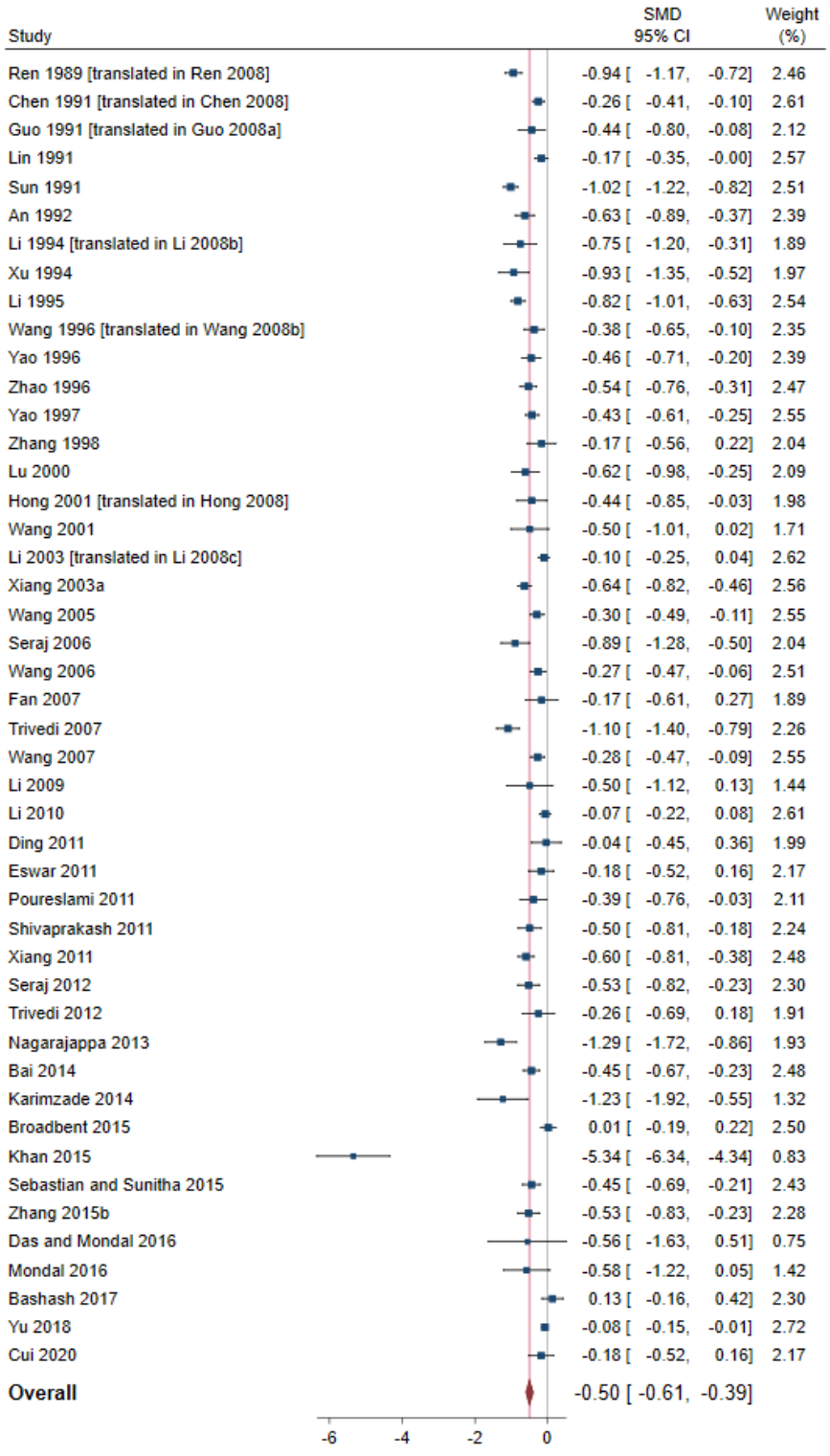


Figure A5-1. Association Between Fluoride Exposure and IQ Scores in Children: Overall Analysis

SMDs for individual studies are shown with solid boxes representing the weight, and the random-effects pooled SMD is shown as a solid diamond. Horizontal lines represent 95% CIs for the study-specific SMDs.

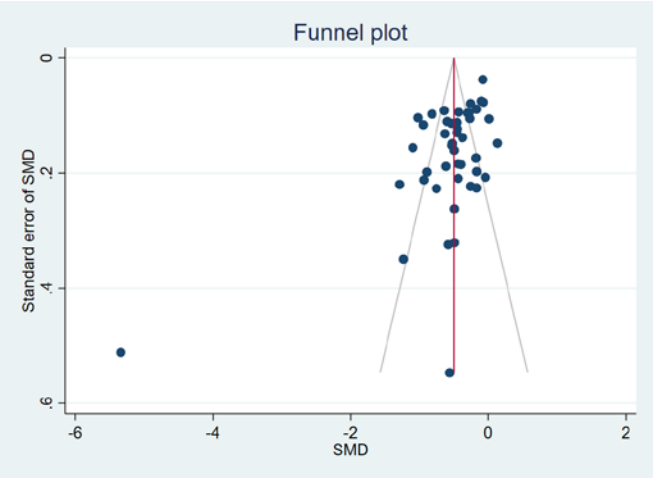


Figure A5-2. Funnel Plot of Included Studies

This funnel plot shows individual studies included in the analysis according to random-effect standardized weighted mean difference (SMD) estimates (x-axis) and the standard error (SE) of each study-specific SMD (y-axis). The solid vertical line indicates the pooled SMD estimate for all studies combined and the dashed lines indicated pseudo 95% confidence limits around the pooled SMD estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -3.21
      SE of beta1 =     0.649
                z =     -4.95
       Prob > |z| =    0.0000


Begg's test for small-study effects

Kendall's score =   -185.00
    SE of score =   105.617
              z =     -1.76
     Prob > |z| =    0.0815
```

Figure A5-3. Test for Publication Bias

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     50
  Model: Random-effects                      observed =     46
 Method: DerSimonian-Laird                   imputed =      4

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.502	-0.611	-0.393
Observed + Imputed	-0.419	-0.537	-0.301

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     56
  Model: Random-effects                      observed =     46
 Method: DerSimonian-Laird                   imputed =      10

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.502	-0.611	-0.393
Observed + Imputed	-0.643	-0.776	-0.511

Figure A5-4. Trim-and-fill Analysis

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator (the linear estimator to the right showed no change in pooled SMD); right panel shows random-effects pooled SMD after filling in to the left using a linear estimator (the run estimator to the left showed no change in pooled SMD).

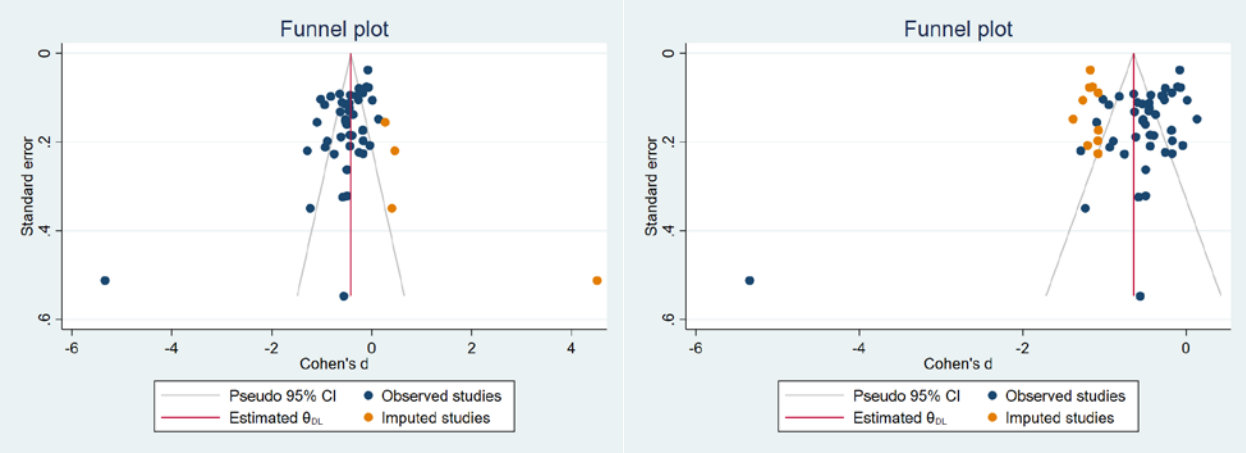


Figure A5-5. Filled-in Funnel Plots to Eliminate Publication Bias

Left panel shows the funnel plot filled in to the right using a run estimator (the linear estimator to the right showed no change in pooled SMD); right panel shows the funnel plot filled in to the left using a linear estimator (the run estimator to the left showed no change in pooled SMD).

Risk-of-bias Subgroup Analysis

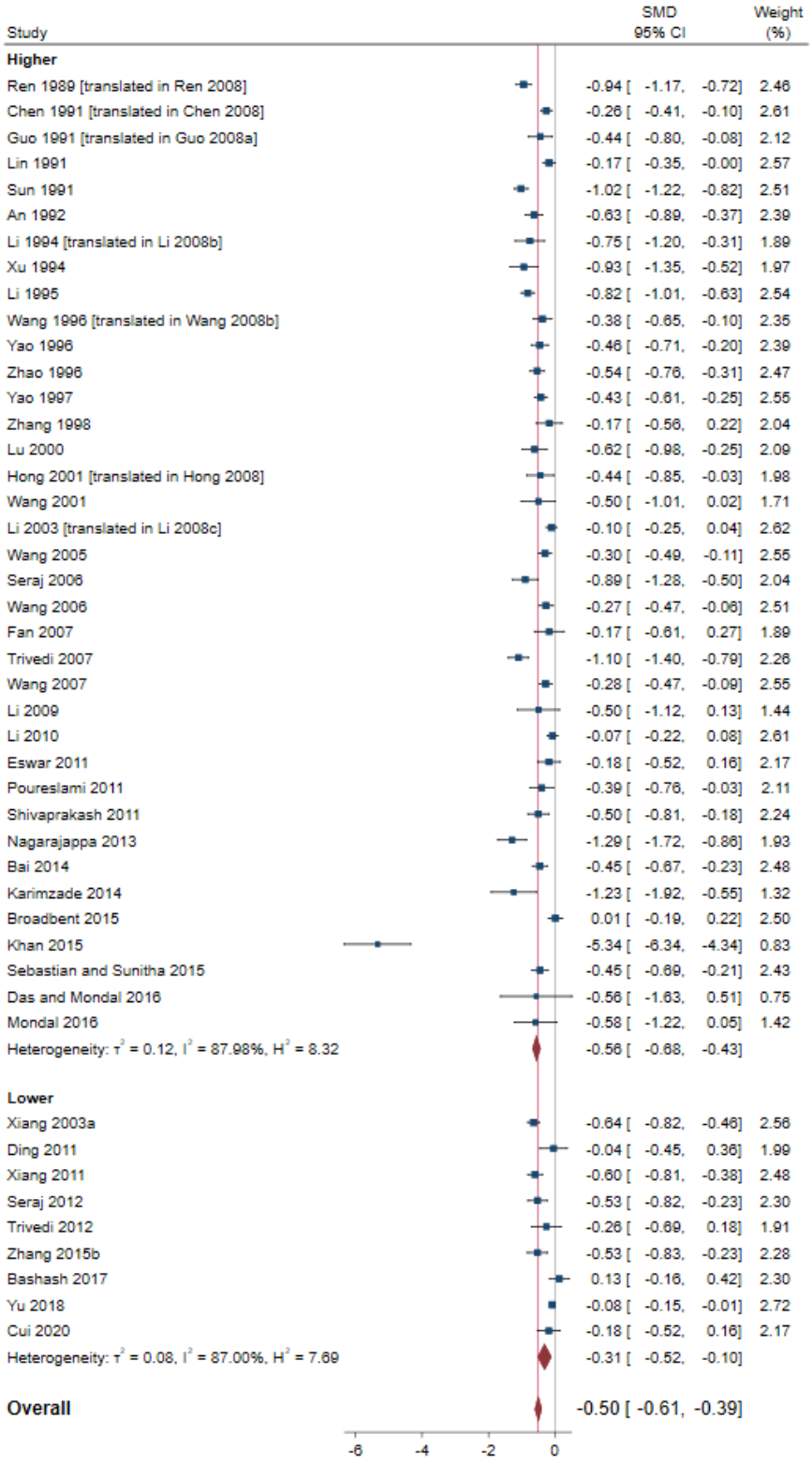


Figure A5-6. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Risk of Bias

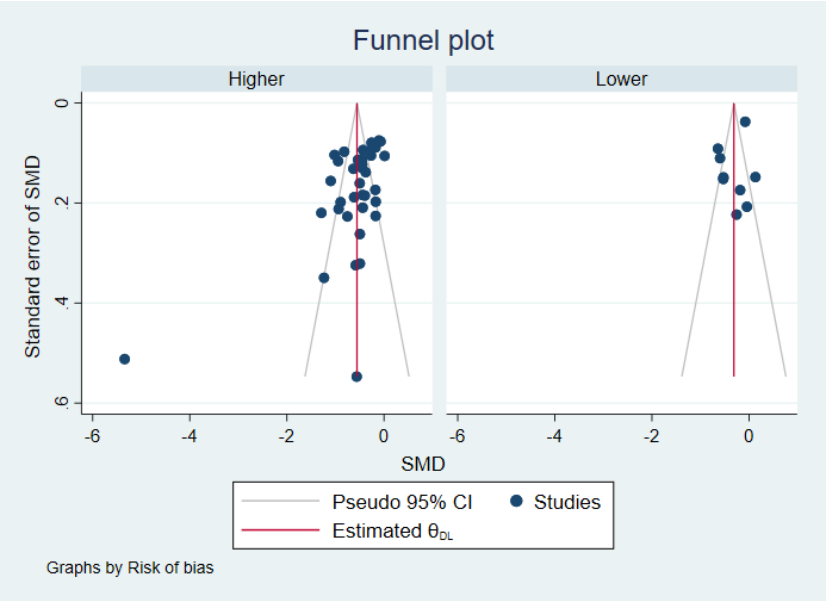


Figure A5-7. Funnel Plot by Risk-of-bias Evaluation

```
Regression-based Egger test for small-study effects
Random-effects model                                  Higher RoB
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -3.63
      SE of beta1 =     0.709
                z =     -5.12
       Prob > |z| =    0.0000

Regression-based Egger test for small-study effects
Random-effects model                                  Lower RoB
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =      0.78
      SE of beta1 =     2.155
                z =      0.36
       Prob > |z| =    0.7169
```

Figure A5-8. Test for Publication Bias by Risk of Bias

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     46
  Model: Random-effects                      observed =     37
 Method: DerSimonian-Laird                   imputed =      9

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.556	-0.684	-0.428
Observed + Imputed	-0.354	-0.498	-0.210

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     44
  Model: Random-effects                      observed =     37
 Method: DerSimonian-Laird                   imputed =      7

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.556	-0.684	-0.428
Observed + Imputed	-0.684	-0.831	-0.537

Figure A5-9. Trim-and-fill Analysis for Higher Risk-of-bias Studies

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator (the linear estimator to the right showed no change in the pooled SMD); right panel shows random-effects pooled SMD after filling in to the left using a linear estimator (the run estimator to the left showed no change in the pooled SMD).

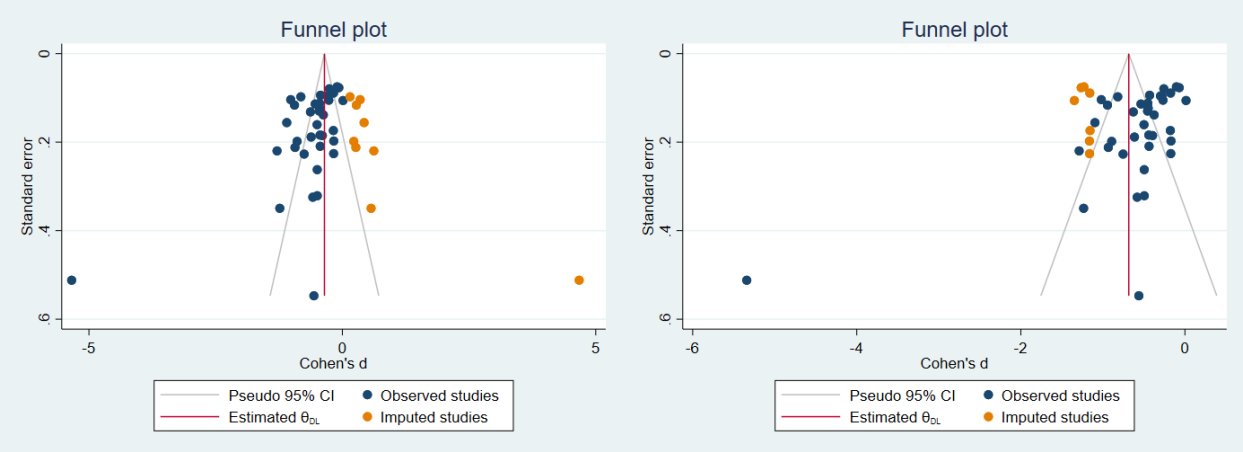


Figure A5-10. Filled-in Funnel Plots to Eliminate Publication Bias for Higher Risk-of-bias Studies

Left panel shows funnel plot filled in to the right using a run estimator (the linear estimator to the right showed no change in the pooled SMD); right panel shows the funnel plot filled in to the left using a linear estimator (the run estimator to the left showed no change in the pooled SMD).

Gender Subgroup Analysis

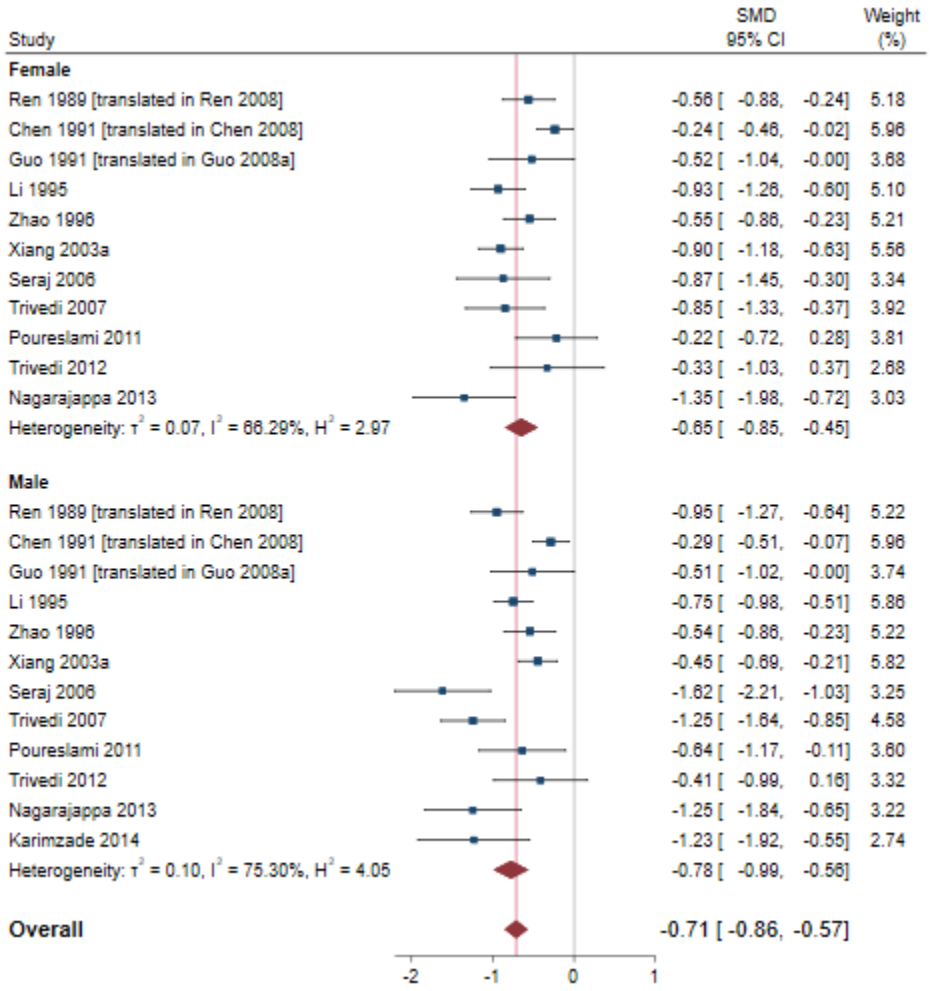


Figure A5-11. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Gender

Age Group Subgroup Analysis

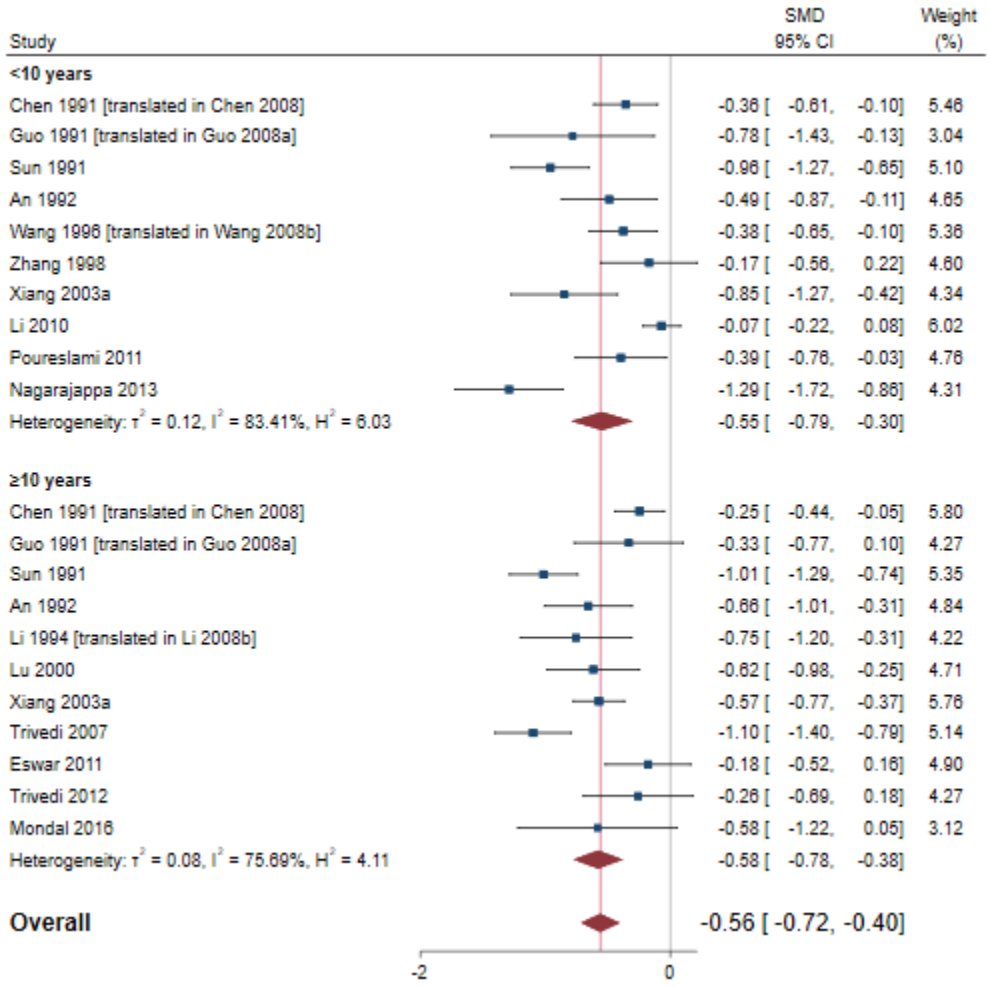


Figure A5-12. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Age Group

Country Subgroup Analysis

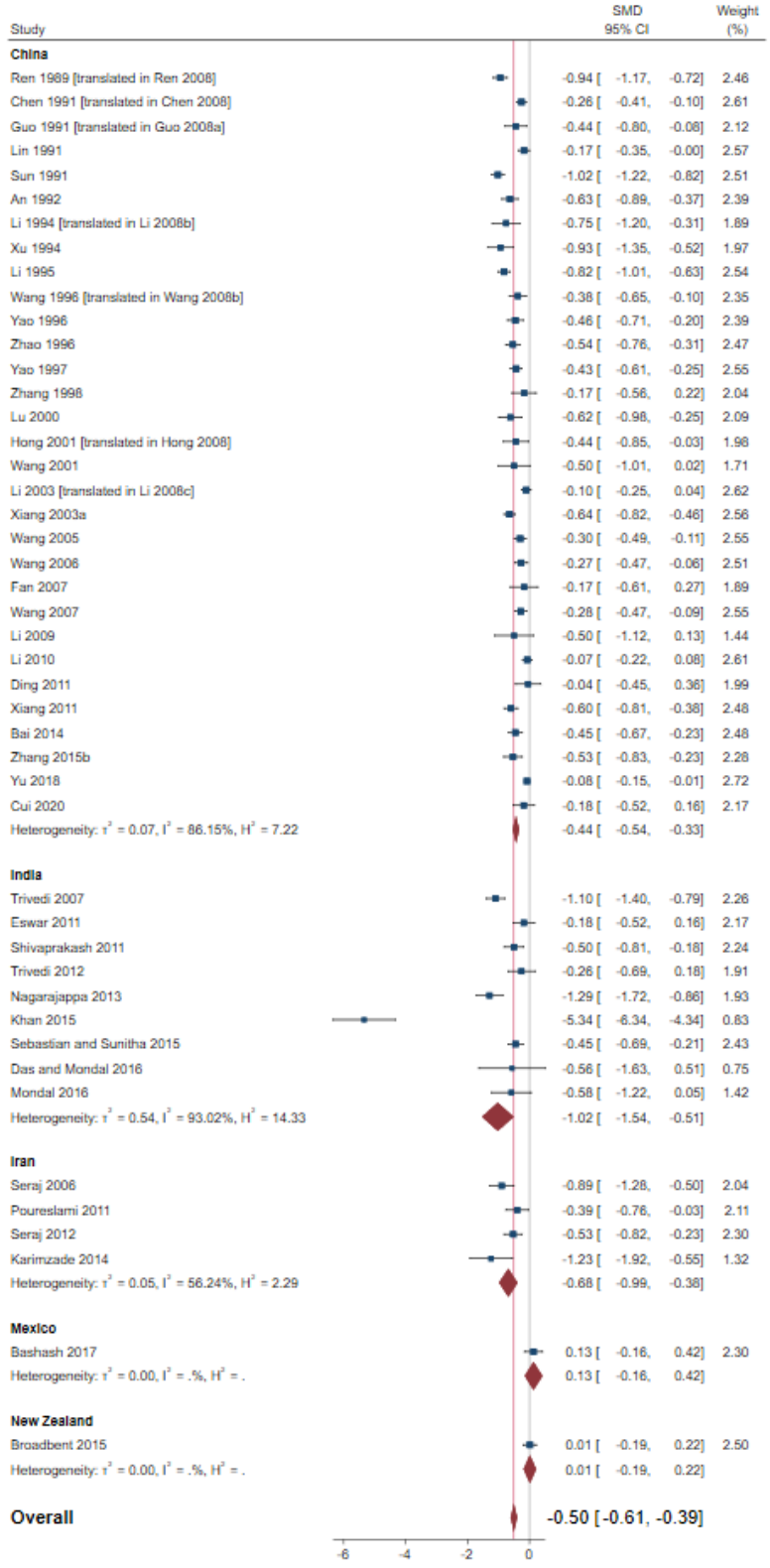


Figure A5-13. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Country

Assessment Type Subgroup Analysis

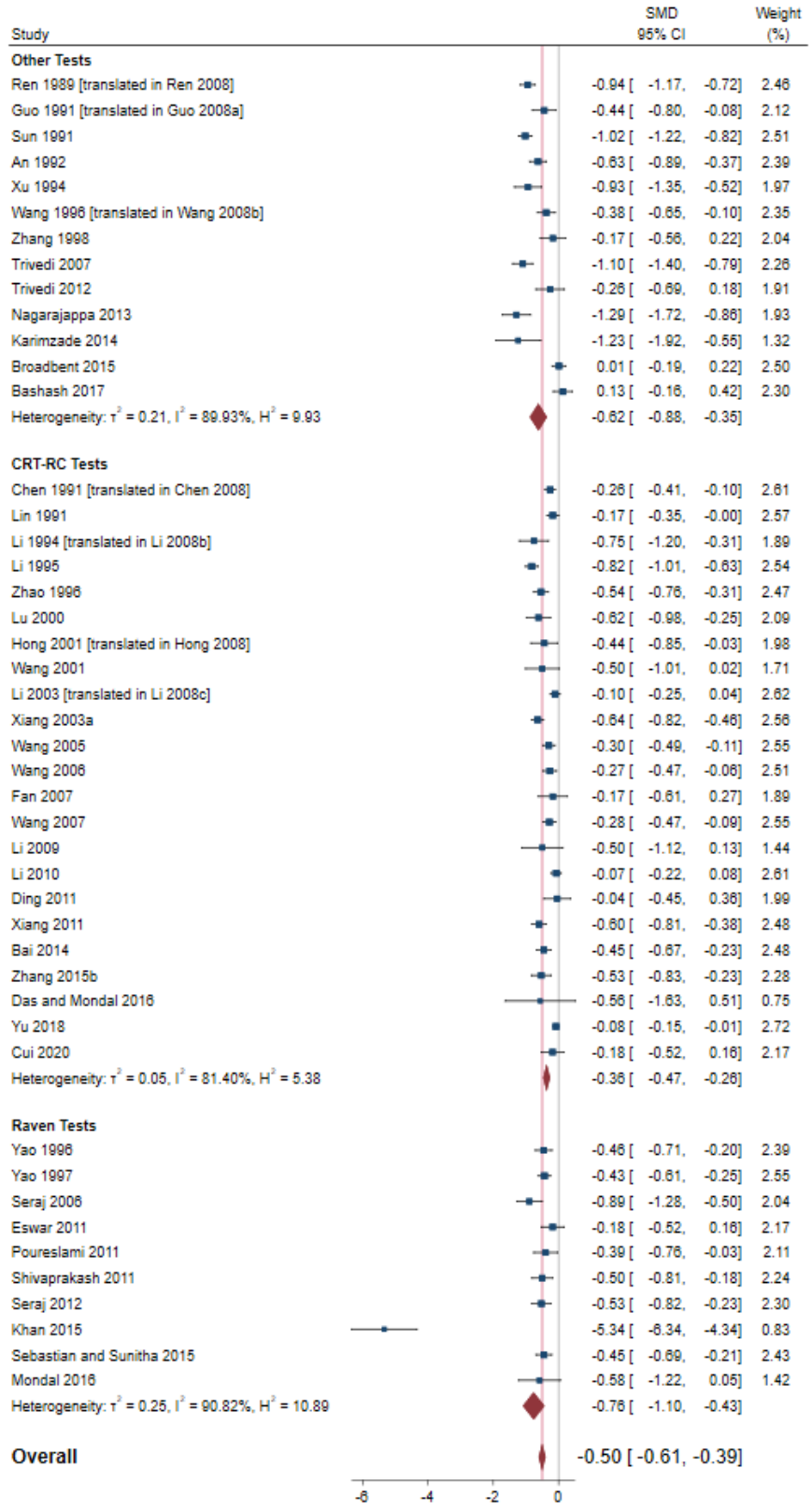


Figure A5-14. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Assessment Type

Exposure Type Subgroup Analysis

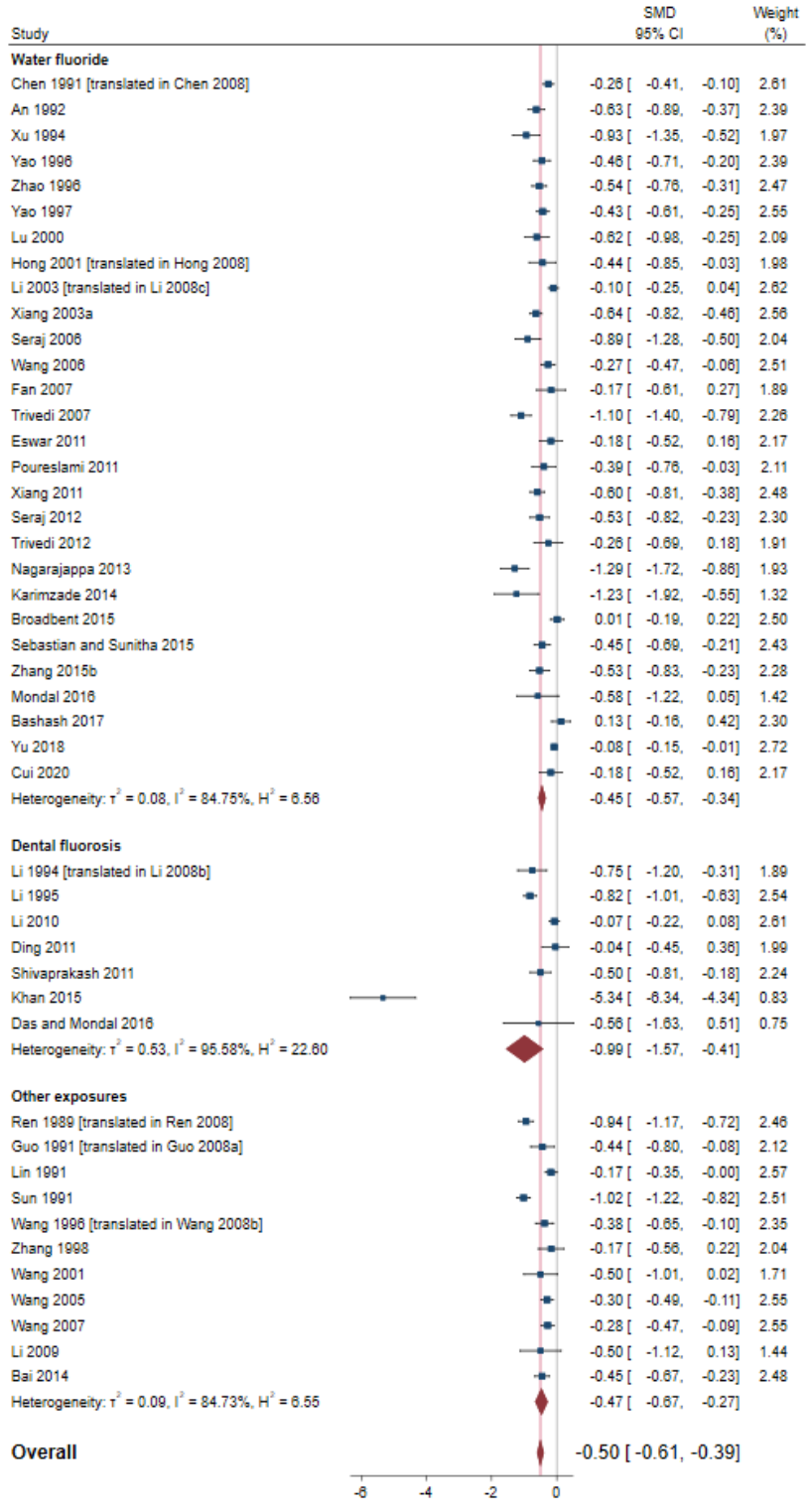


Figure A5-15. Association Between Fluoride Exposure and IQ Scores in Children: Effect by Exposure Type

Exposure types include water, dental fluorosis, and other exposures (iodine, arsenic, aluminum, and fluoride from coal burning).

Aim 2. To conduct new meta-analyses using individual-level exposure data

Approach

The individual-effect meta-analysis included studies that reported effect estimates as betas and included 95% CIs or SEs (see **Table B-1**). Adjusted effect estimates were used in the meta-analysis. If results from multiple models were reported within a single study, the most adjusted results were selected. The study outcomes were evaluated with respect to a 1-mg/L unit increase in exposure. To ensure consistent units across studies, units of exposure were transformed to mg/L as needed. For Bashash *et al.* (2017), Yu *et al.* (2018), and Till *et al.* (2020), units of exposure were transformed to levels ranging from 0.5 to 1 mg/L. For Cui *et al.* (2018), units of exposure were transformed from 1 log mg/L to 1 mg/L. Cui *et al.* (2018) reported an association between IQ and log-transformed exposure. A sensitivity analysis was performed to evaluate the impact of using Cui *et al.* (2018), since the relationship between IQ and exposure evaluated in this study was not linear (as in the other studies included). Yu *et al.* (2018) reported estimates from piecewise linear regression models, with three estimated ranges for urinary fluoride exposure (low 0.01−1.60 mg/L, medium 1.60−2.50 mg/L, and high 2.50−5.54 mg/L) and two estimated ranges for water fluoride exposure (low 0.20−3.40 mg/L and high 3.40–3.90 mg/L). Because these piecewise effect estimates are likely correlated, study-specific pooled effect estimates were used for urinary and water fluoride exposures for the overall effect meta-analysis. A sensitivity analysis was performed to evaluate the impact of using the pooled estimate rather than the piecewise estimates from Yu *et al.* (2018).

Yu *et al.* (2018) and Wang *et al.* (2020b) used the same study cohort of children recruited in 2015 from the rural areas of Tianjin City, China. Only results from Yu *et al.* (2018) were included in the meta-analysis since Wang *et al.* (2020b) used a subset (n = 571) of the original study population from Yu *et al.* (2018) (n = 2,668). A sensitivity analysis was conducted to evaluate the impact of using the effect estimate from Wang *et al.* (2020b) rather than the pooled effect estimate from Yu *et al.* (2018).

Green *et al.* (2019) and Till *et al.* (2020) used the same cohort of 398 mother-child dyads in the Maternal-Infant Research on Environmental Chemicals (MIREC) cohort that reported drinking tap water in 10 Canadian cities. Both studies reported effect estimates for maternal urinary fluoride (MUF) and water fluoride concentrations. In the Green *et al.* (2019) study, 512 mother-child pairs had MUF data (and all covariates) compared to 398 pairs in Till *et al.* (2020). Water fluoride levels were available for 420 pairs in Green *et al.* (2019) compared to 398 pairs in Till *et al.* (2020). Both references reported effect estimates adjusted for maternal education, maternal race, child's sex, HOME total score, and secondhand smoking status in the child's residence. In addition, Till *et al.* (2020) adjusted for child's age at IQ testing. The age range for study subjects was 3–4 years old. For the main analysis, results from Green *et al.* (2019) were included. A sensitivity analysis was conducted using the water fluoride result for formula-fed children and the MUF result from Till *et al.* (2020) since these are most adjusted compared to Green *et al.* (2019). For fluoride intakes, estimates from both studies were used including total fluoride intake from Green *et al.* (2019) and infant formula intake from Till *et al.* (2020).

In the overall effect analysis, for studies reporting multiple measures of fluoride exposure, results associated with measured or estimated individual-level exposures, biomarker levels (such as urinary fluoride), or fluoride intake were prioritized over results associated with water fluoride concentrations. Subgroup analyses were performed that considered all exposure types. All studies used in these analyses with individual-level effects were lower risk of bias.

The overall effect based on studies with individual-level measures of exposure showed that a 1-unit increase in fluoride exposure was associated with a significantly lower IQ score (beta = −1.40; 95% CI: −2.33, −0.47) (**Table A5-2**). A 1-mg/day increase in fluoride intake resulted in significantly lower IQ score (beta = −3.31; 95% CI: −6.12, −0.50). A 1-mg/L increase in water fluoride also resulted in significantly lower IQ score (beta = −4.77; 95% CI: −9.10, −0.45). The results for fluoride intake and water fluoride, however, are based on two studies and should be interpreted with caution.

There was evidence of moderate heterogeneity (I^2 = 46%, p = 0.101; **Table A5-2**, **Figure A5-16,** and **Figure A5-17**) in studies with individual-level urinary exposure levels. Eliminating publication bias through trim-and-fill analysis continued to support that 1-mg/L increases in individual-level urinary or water fluoride were associated with lower IQ scores, with an adjusted pooled effect estimate of −0.82 (95% CI: −1.81, 0.17) (**Figure A5-19**).

A sensitivity analysis to evaluate the impact of using the piecewise estimates from Yu *et al.* (2018) revealed no significant change in the pooled effect estimate (−1.37; 95% CI: −2.38, −0.37) (**Figure B-1**). A sensitivity analysis to evaluate using estimates from Wang *et al.* (2020b) rather than Yu *et al.* (2018) study also revealed no significant change in the pooled effect estimate (−1.24; 95% CI: −1.94, −0.54) (**Figure B-6**). A sensitivity analysis using the water fluoride result for formula-fed children and the MUF result from Till *et al.* (2020) (rather than Green *et al.* (2019)) suggested no significant change in the overall pooled effect estimate (−1.50; 95% CI: −2.44, −0.57) (**Figure B-11**).

Summary Results

Table A5-2. Pooled Effect Estimates and 95% CIs for Children's IQ Scores and Individual-level Exposures to Fluoride

Analysis	Number of Studies	Beta (95% CI)	Heterogeneity	
			p-value	I^2
Overall Effect				
Full-scale IQ	6	−1.40 (−2.33, −0.47)	0.101	46%
Verbal IQ	6	−1.36 (−2.28, −0.45)	0.110	44%
Performance IQ	6	−1.33 (−2.25, −0.42)	0.117	43%
Subgroup Analyses				
Gender				
Males	2	−2.23 (−5.45, 0.99)	0.092	65%
Females	3	−1.64 (−4.80, 1.51)	0.045	68%
Country				
Canada	1	−1.95 (−5.18, 1.29)	NA	
China	4	−0.84 (−1.39, −0.30)	0.342	10%
Mexico	1	−5.00 (−8.53, −1.47)	NA	
Assessment Type				
CRT-RC tests	4	−0.84 (−1.39, −0.30)	0.342	10%
Non-CRT-RC tests	2	−3.39 (−6.37, −0.41)	0.212	36%
Exposure Type				
Urinary fluoride	6	−1.40 (−2.33, −0.47)	0.101	46%
Intake	2	−3.31 (−6.12, −0.50)	0.746	0%
Water fluoride	2	−4.77 (−9.09, −0.45)	0.707	0%

Notes:
CI = confidence interval; NA = not applicable

Overall Analysis

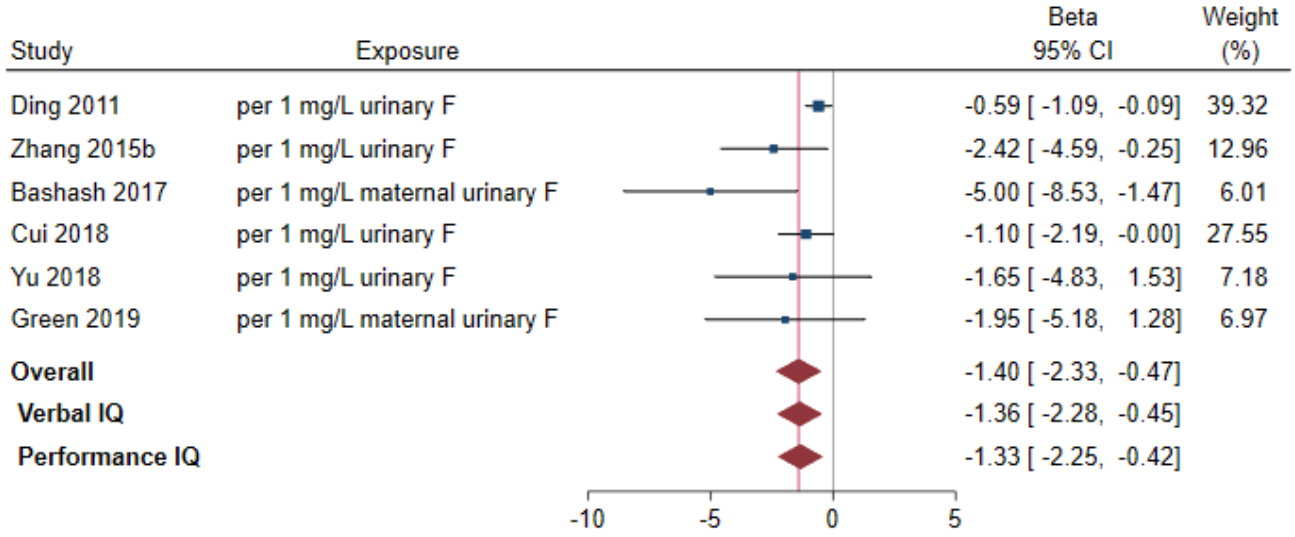


Figure A5-16. Association Between Individual-level Fluoride Exposure and IQ Scores in Children: Overall Analysis

Estimates (betas) for individual studies are shown with solid boxes representing the weight, and the pooled estimate is shown as a solid diamond. Horizontal lines represent 95% Cis for the study-specific betas.

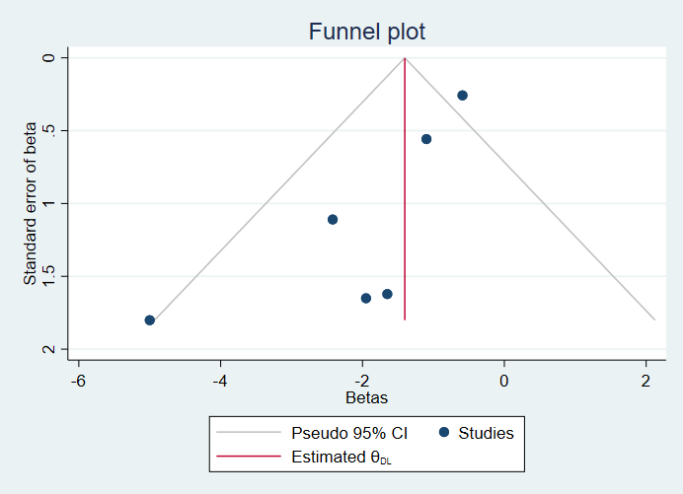


Figure A5-17. Funnel Plot of Included Studies with Individual-level Exposures

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.68
      SE of beta1 =     0.628
                z =     -2.67
       Prob > |z| =    0.0075
```

Figure A5-18. Test for Publication Bias for Studies with Individual-level Exposures

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration                           Number of studies =      9
  Model: Random-effects                      observed =      6
 Method: DerSimonian-Laird                   imputed =       3

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Beta	[95% Conf.	Interval]
Observed	-1.403	-2.333	-0.473
Observed + Imputed	-0.818	-1.809	0.172

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     10
  Model: Random-effects                      observed =      6
 Method: DerSimonian-Laird                   imputed =       4

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Beta	[95% Conf.	Interval]
Observed	-1.403	-2.333	-0.473
Observed + Imputed	-0.737	-1.671	0.197

Figure A5-19. Trim-and-fill Analysis for Studies with Individual-level Exposures

Left panel shows the random-effects pooled effect estimate after filling in to the right using a liner estimator; right panel shows random-effects pooled effect estimate after filling in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled slope.

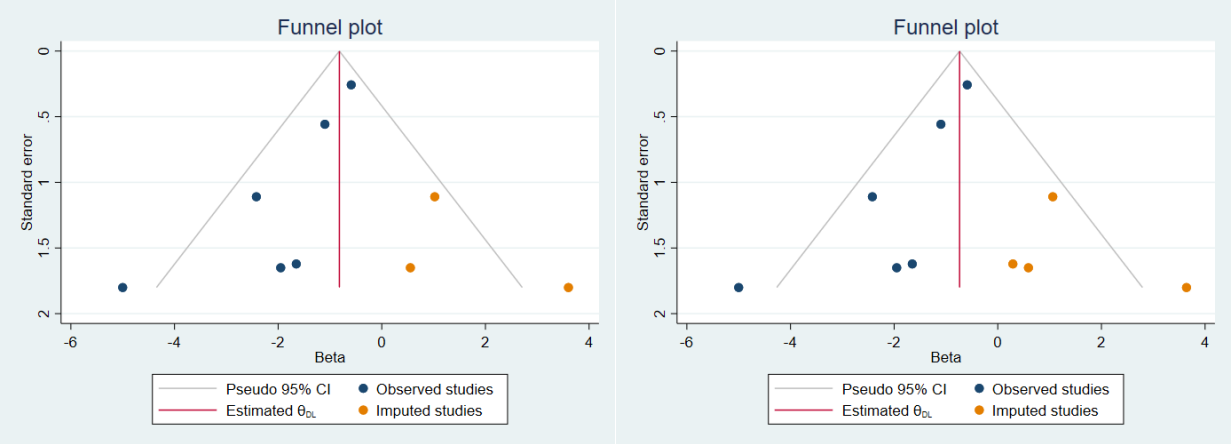


Figure A5-20. Filled-in Funnel Plots to Eliminate Publication Bias for Studies with Individual-level Exposures

Left panel shows funnel plot filled in to the right using a linear estimator; right panel shows the funnel plot filled in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

Exposure Type Subgroup Analysis

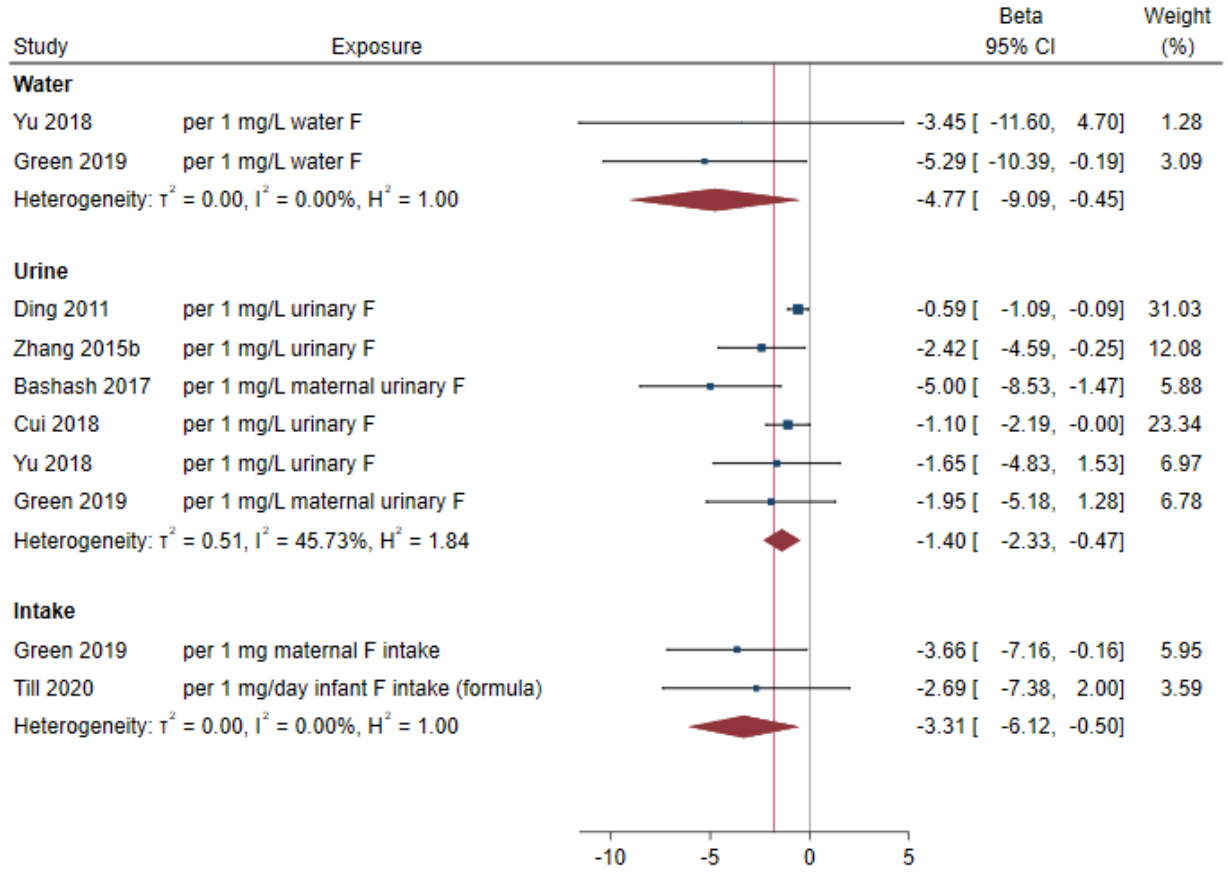


Figure A5-21. Association Between Individual-level Fluoride Exposure and IQ Scores in Children: Effect by Exposure Type

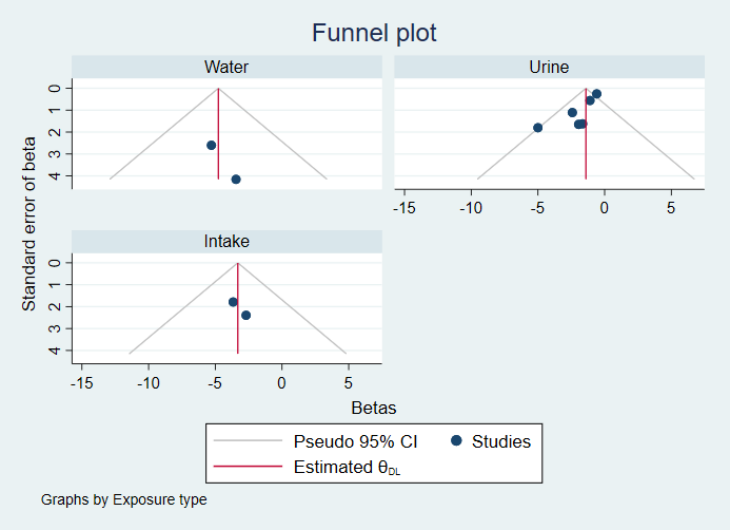


Figure A5-22. Funnel Plot of Included Studies

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analyses. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Country Subgroup Analysis

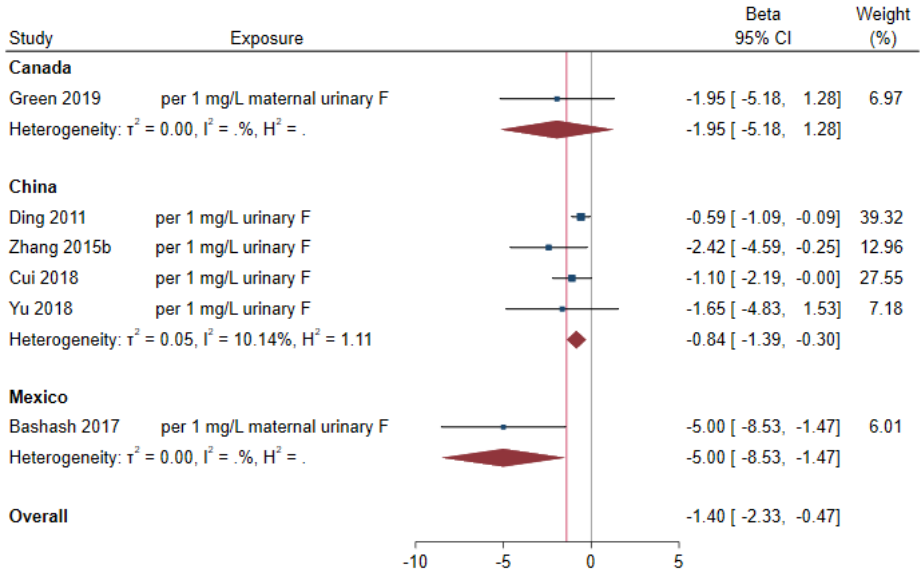


Figure A5-23. Association Between Individual-level Fluoride Exposure and IQ Scores in Children: Effect by Country

Note: The analyses for publication bias for studies from China, Canada, and Mexico rely on a very small number of studies each and are not shown.

Assessment Type Subgroup Analysis

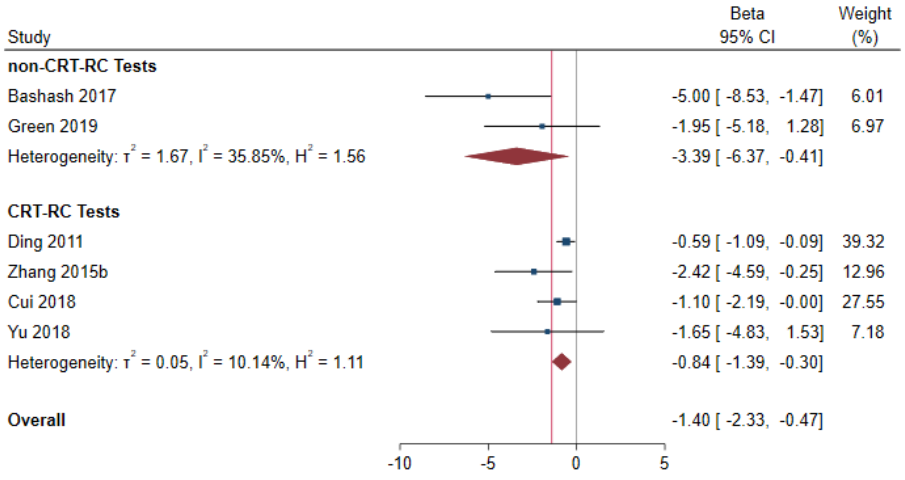


Figure A5-24. Association Between Individual-level Fluoride Exposure and IQ Scores in Children: Effect by Assessment Type

Note: The analyses for publication bias for CRT-RC studies and non-CRT-RC studies include only four and two studies, respectively, and are not shown.

Gender Subgroup Analysis

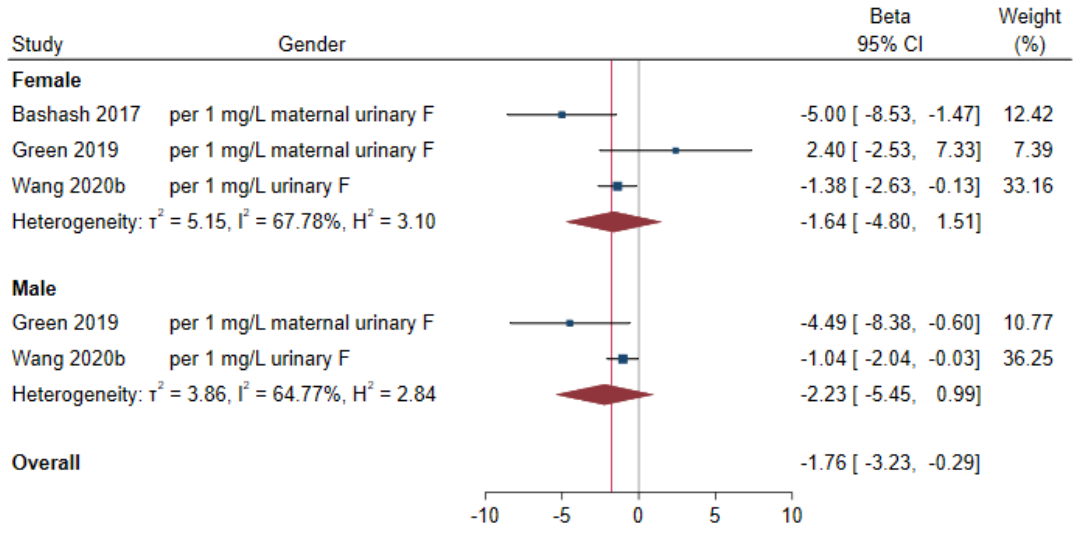


Figure A5-25. Association Between Individual-level Fluoride Exposure and IQ Scores in Children: Effect by Gender

Note: The analysis for publication bias by gender relies on two or three studies each and are not shown.

Dose-response meta-analyses using mean-effect estimates

Approach

For the dose-response meta-analysis using the mean-effect estimates, a one-step approach developed in the protocol (https://ntp.niehs.nih.gov/go/785076) and Crippa *et al.* (2018) was used. The approach uses linear mixed models to analyze all available data, including studies with one nonreference group. For each study, the median or mean fluoride intake for each exposure group was assigned to its corresponding effect estimate. If median or mean intakes by exposure group were not provided, the midpoint of the upper and lower boundaries in every category was assigned as the average intake. If the upper boundary for the highest exposure group was not reported, the boundary was assumed to have the same amplitude as the nearest category. For each study, the SMDs and corresponding SEs are used to compare the differences in mean IQ between the exposed and reference groups. The corresponding SMD for the reference group is set to zero for this analysis. The SMDs and corresponding variances are used to estimate a pooled dose-response curve using a restricted maximum likelihood estimation method. To examine a potential nonlinear relationship between exposure to fluoride and children's IQ levels, quadratic terms and restricted cubic splines were created, and a potential departure from a linear trend was assessed by testing the coefficient of the quadratic term and a second spline equal to zero. Models were compared and best model fit was determined based on the Akaike information criterion (AIC) (Müller *et al.* 2013).

A dose-response meta-analysis using the effect estimates reported in studies with individual-level exposure was considered. However, because of the small number of studies (n = 10), the various types of exposure metrics, and the different types of reported effect estimates that could not be combined, a dose-response meta-analysis of these studies could not be conducted.

Summary Results

Characteristics of the studies that compared mean IQ scores among groups of children with different levels of fluoride exposure are shown in **Table A-1**.

The meta-analysis combining data from 31 studies with fluoride levels in drinking water showed a significantly lower children's IQ score with increasing exposure (**Table A5-3** A5-3). Based on the AIC, the best model fit was achieved when restricted cubic splines with three knots were added to the linear model. However, given the small difference in AICs between the different models, and for ease of interpretability, the linear model results were chosen for the purposes of discussion although results from all models are presented in **Table A5-3.** Based on the linear model, the decrease in mean SMD between exposed and reference groups was −0.14 (95% CI: −0.19, −0.08) (**Table A5-3**). When the analysis was restricted to studies with the "high" group exposed to < 1.5 mg/L fluoride in drinking water (n = 9; 2 lower risk-of-bias studies and 7 higher risk-of-bias studies), the mean SMD became positive and nonsignificant (0.32; 95% CI: −0.57, 1.20). However, when including groups exposed to < 2 mg/L fluoride in drinking water, the mean SMD in children's IQ scores was both negative and statistically significant (SMD = −0.27; 95% CI: −0.36, −0.17) (n = 9; 2 lower risk-of-bias studies and 7 higher risk-of-bias studies).

The meta-analysis combining data from 9 studies with fluoride levels in urine showed a significantly lower children's IQ with increasing exposure (SMD = −0.18; 95% CI: −0.31, −0.05) (**Table A5-3**). Based on AIC, the best model fit was the linear model (**Table A5-3**). There was no improvement in the fit of the model when a quadratic term or restricted cubic splines were added to the linear model (**Table A5-3**). When the analyses were restricted to studies with the "high" group with < 1.5 mg/L fluoride in urine (n = 4; 2 lower risk-of-bias studies and 2 higher risk-of-bias studies), the direction of the effect did not

change, but it was no longer statistically significant (SMD = −0.13; 95% CI: −0.29, 0.03). When the dose-response meta-analysis was extended to include exposed groups with < 2 mg/L urinary fluoride (n = 6; 3 lower risk-of-bias studies and 3 higher risk-of-bias studies), the mean SMD remained negative and not statistically significant (−0.25; 95% CI: −0.71, 0.22).

Dose-Response Meta-analysis Using Mean Effects—Model Selection

Table A5-3. Model Comparison for Dose-response Meta-analysis for Children's IQ Scores (SMDs) and Exposures to Fluoride: Parameter Estimates and Model Fit[1]

Analysis	No. of Studies/ No. of Observations	Linear Model[2]	Quadratic Model[3]	Restricted Cubic Splines Model[4]
Water Fluoride				
All data	31/49	−0.14 (−0.19, −0.08) 113.6	−0.23 (−0.32, −0.14) 0.02 (0.01, 0.03) 110	−0.24 (−0.40, −0.08) 0.38 (−0.06, 0.81) 101.7
<1.5 mg/L	9/12	0.32 (−0.57, 1.20) 26.1	1.97 (−0.98, 4.92) −1.26 (−3.23, 0.71) 28.8	0.81 (−0.37, 1.99) −19.37 (−42.19, 3.44) 22.4
<2 mg/L	9/17	−0.27 (−0.36, −0.17) 44.6	0.64 (−1.04, 2.32) −0.40 (−1.09, 0.29) 31.6	0.23 (−0.71, 1.17) −4.63 (−12.47, 3.21) 27.6
<4 mg/L	24/38	−0.17 (−0.25, −0.09) 81.7	−0.28 (−0.64, 0.08) 0.03 (−0.10, 0.16) 79.5	−0.26 (−0.48, −0.03) 0.36 (−1.06. 1.79) 73.9
Urinary Fluoride				
All data	9/22	−0.18 (−0.31, −0.05) 61.1	−0.17 (−0.46, 0.12) −0.004 (−0.05, 0.04) 71.6	−0.12 (−0.31, 0.07) −0.11 (−0.35, 0.13)) 68.7
<1.5 mg/L	4/7	−0.13 (−0.29, 0.03) −0.3	−0.63 (−1.28, 0.01) 0.31 (−0.06, 0.68) 4.7	−0.30 (−0.55, −0.04) 5.09 (−1.00, 11.18) −0.9
<2 mg/L	6/11	−0.09 (−0.22, 0.03) −2.3	−0.25 (−0.71, 0.22) 0.07 (−0.13, 0.27) 5.8	−0.13 (−0.34, 0.09) 0.25 (−0.87, 1.37) 2.9

Notes:
AIC = Akaike information criterion; SMD = standardized mean difference
[1]Parameter estimates are changes in SMDs (beta [95% CI]); model fit is represented by the AIC.
[2]The estimates represent change in SMD for the linear model and AIC, respectively.
[3]The estimates represent change in SMD for the linear term, change in SMD for quadratic term, and AIC, respectively.
[4]The estimates represent change in SMD for the first spline term, change in SMD for the second spline term, and AIC, respectively.

Attachment A. Subgroup and Sensitivity Analyses (Aim 1)

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Ren *et al.* (1989) [translated in Ren *et al.* 2008]	China	8–14	No fluoride measurement High fluoride and low iodine village/low iodine village	Not specified	Wechsler Intelligence Scale for Children	Higher	Gender; iodine
Chen *et al.* (1991) [translated in Chen *et al.* 2008][w]	China	7–14	Drinking water Endemic fluorosis village/nonendemic village	0.89 mg/L (nonendemic); 4.55 mg/L (endemic)	Chinese Standardized Raven Test	Higher	Age; gender
Guo *et al.* (1991) [translated in Guo *et al.* 2008a]	China	7–13	Serum Coal burning-related fluoride endemic area/control area using wood	0.1044 ± 0.0652 mg/L (control); 0.1483 ± 0.0473 mg/L (endemic)	Chinese Binet Intelligence Test	Higher	Age; gender; SES
Lin *et al.* (1991) [w*]	China	7–14	Drinking water High fluoride and low iodine village/low iodine village/control area with iodine supplementation	0.34 mg/L (low iodine village); 0.88 mg/L (high fluoride, low iodine village)	Combined Raven's Test for Rural China	Higher	SES
Sun *et al.* (1991)	China	6.5–12	No fluoride measurement Endemic (aluminum-fluoride endemic toxicosis)/nonendemic	Fluorosis: 98.36% (endemic)	Japan's Shigeo Kobayashi's 50-point scoring method	Higher	Age
An *et al.* (1992)[w]	China	7–16	Drinking water High fluoride/nonhigh fluoride area	0.6−1.0 mg/L (nonhigh); 2.1−7.6 mg/L (high)	Wechsler Intelligence Scale for Children-Revised	Higher	Age; race; SES
Li *et al.* (1994) [translated in Li *et al.* 2008b]	China	12–13	Grain (cooked by burning high-fluoride coal) High fluoride group III (dental fluorosis present)/high fluoride group II (dental fluorosis present)/high fluoride group I (no dental fluorosis)/control group (no dental fluorosis)	0.5 mg/kg (reference); 4.7 mg/kg (group I); 5.2 mg/kg (group II); 31.6 mg/kg (group III)	Proofing test	Higher	Age; gender; SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Xu *et al.* (1994)[w*]	China	8–14	Drinking water Low- and high-fluoride and iodine regions/control region	0.8 mg/L (reference); 0.8 mg/L (low iodine); 0.5 mg/L (low fluoride, low iodine); 0.38 mg/L (low fluoride); 0.5 mg/L (low fluoride, high iodine); 2.0 mg/L (high fluoride, low iodine); 1.8 mg/L (high fluoride); 3.9 mg/L (high fluoride, high iodine);	Binet-Simon Scale	Higher	–
Li *et al.* (1995)[u]	China	8–13	Urine Dental fluorosis index (DFI) Fluorosis area due to soot from coal burning/nonfluorosis area	1.02 mg/L; DFI: <0.4 (nonfluorosis area); 1.81 mg/L; DFI: 0.8 (slight fluorosis area); 2.01 mg/L; DFI: 2.5 (medium fluorosis area); 2.69 mg/L; DFI: 3.2 (severe fluorosis area)	China Rui Wen Scaler for Rural Areas	Higher	Gender
Wang *et al.* (1996) [translated in Wang *et al.* 2008b][w]	China	4–7	Drinking water (well) High fluoride region/low fluoride region	0.58−1.0 mg/L (low); >1.0−8.6 mg/L (high)	Wechsler Preschool and Primary Scale of Intelligence	Higher	Age; gender
Yao *et al.* (1996)[w]	China	8–12	Drinking water Endemic fluorosis area/nonendemic area	1 mg/L (nonendemic); 2 mg/L (slightly endemic); 11 mg/L (severely endemic)	Raven Test – Associative Atlas	Higher	Iodine; SES
Zhao *et al.* (1996)[w]	China	7–14	Drinking water High fluoride village (Sima)/low fluoride village (Xinghua)	0.91 mg/L (low) 4.12 mg/L (high)	China Rui Wen Scaler for Rural Areas	Higher	Age; SES
Yao (1997)[w*]	China	7–12	Drinking water Fluorosis area without water improvements/fluorosis area with water improvements/nonfluorosis area	0.4 mg/L (nonfluorosis area); 0.33 mg/L (fluorosis area with water improvement); 2 mg/L (fluorosis area without water improvement)	Raven's Standard Progressive Matrices (China's Rural Version)	Higher	Iodine; SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Zhang *et al.* (1998)[w*]	China	4–10	Drinking water High fluoride and high arsenic group/high fluoride group/reference	0.58 mg/L (reference); 0.8 mg/L (high fluoride, high arsenic); 0.8 mg/L (high fluoride)	Shigeo Kobayashi 50-pt. test	Higher	Age; arsenic
Lu *et al.* (2000)[w,u]	China	10–12	Urine, drinking water High fluoride area/low fluoride area	0.37 ± 0.04 mg/L drinking water (low); 1.43 ± 0.64 mg/L urine (low); 3.15 ± 0.61 mg/L drinking water (high); 4.99 ± 2.57 mg/L urine (high)	Chinese Combined Raven Test-C2	Higher	SES
Hong *et al.* (2001) [translated in Hong *et al.* 2008][w*]	China	8–14	Drinking water High fluoride and iodine regions/reference	0.75 mg/L (reference); 2.90 mg/L (high fluoride); 2.85 mg/L (high fluoride, high iodine); 2.94 mg/L (high fluoride, low iodine); 0.48 mg/L (low fluoride, low iodine)	Chinese Standardized Raven Test	Higher	Iodine; SES; demographics
Wang *et al.* (2001)[w]	China	8–12	Drinking water Investigative point (high fluoride)/control point (low fluoride)	0.5 mg/L (low); 2.97 mg/L (high)	Combined Raven's Test for Rural China	Higher	–
Li *et al.* (2003) [translated in Li *et al.* 2008c]	China	6–13	No fluoride measurement Endemic fluorosis areas/reference	Not specified	Chinese Standardized Raven Test	Higher	–

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Xiang *et al.* (2003a)[w*]	China	8–13	Drinking water Endemic fluorosis areas/nonendemic fluorosis area	0.36 ± 0.15 mg/L (reference); 0.75 ± 0.14 mg/L (endemic fluorosis area group A); 1.53 ± 0.27 mg/L (endemic fluorosis area group B); 2.46 ± 0.3 mg/L (endemic fluorosis area group C); 3.28 ± 0.25 mg/L (endemic fluorosis area group D); 4.16 ± 0.22 mg/L (endemic fluorosis area group E); 2.47 ± 0.79 mg/L (high fluoride)	Combined Raven's Test for Rural China	Lower	Age; gender; iodine; lead; SES
Wang *et al.* (2005)	China	8–12	Drinking water High fluoride group/reference	0.48 mg/L (reference); 8.31 mg/L (high)	Chinese Combined Raven Test-C2	Higher	SES
Seraj *et al.* (2006)[w]	Iran	7–11	Drinking water High fluoride area/low fluoride area	0.4 ppm (low); 2.5 ppm (high area)	Raven Test	Higher	Gender
Wang *et al.* (2006)[w]	China	8–12	Drinking water Area severely affected by fluorosis/reference	0.73 ± 0.28 mg/L (reference); 5.54 ± 3.88 mg/L (high)	Combined Raven's Test for Rural China	Higher	–
Fan *et al.* (2007)[w]	China	7–14	Drinking water High fluoride area/low fluoride area	1.03 mg/L (low); 3.15 mg/L (high)	Chinese Combined Raven Test-C2	Higher	–
Trivedi *et al.* (2007)[w]	India	12–13	Drinking water High fluoride area/low fluoride area	2.01 ± 0.009 mg/L (low); 5.55 ± 0.41 mg/L (high)	questionnaire prepared by Professor JH Shah	Higher	Age; gender
Wang *et al.* (2007)[w]	China	8–12	Drinking water High fluoride area/reference	0.5 ± 0.2 mg/L (reference); 8.3 ± 1.9 mg/L (high)	Combined Raven's Test for Rural China	Higher	Age; gender; arsenic; SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Li *et al.* (2009)[u*]	China	8–12	Urine Endemic fluorosis region caused by coal burning (severe/medium/mild/reference) Degree of dental fluorosis (severe/medium/mild/very mild/suspected/normal)	0.962 ± 0.517 mg/L (reference) 1.235 ± 0.426 mg/L (mild endemic region) 1.670 ± 0.663 mg/L (medium endemic region) 2.336 ± 1.128 mg/L (severe endemic region) 0.867 ± 0.233 mg/L (normal fluorosis) 1.094 ± 0.355 mg/L (suspected fluorosis) 1.173 ± 0.480 mg/L (very mild fluorosis) 1.637 ± 0.682 mg/L (mild fluorosis) 2.005 ± 0.796 mg/L (medium fluorosis) 2.662 ± 1.093 mg/L (severe fluorosis)	Combined Raven's Test for Rural China	Higher	Age; gender
Li *et al.* (2010)	China	7–10	No fluoride measurement Dental fluorosis children/nondental fluorosis children	Not specified	Combined Raven's Test for Rural China	Higher	Gender

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Ding *et al.* (2011)[u*]	China	7–14	Urine Dental fluorosis (questionable/moderate/mild/very mild/normal) Mean urinary fluoride levels (10 groups)	0.80 ± 0.55 mg/L (normal dental fluorosis); 1.11 ± 0.74 mg/L (very mild dental fluorosis); 1.31 ± 0.78 mg/L (mild dental fluorosis); 1.46 ± 0.79 mg/L (moderate dental fluorosis); 1.13 ± 0.73 mg/L (questionable dental fluorosis); 0.26 mg/L (group 1); 0.45 mg/L (group 2); 0.56 mg/L (group 3); 0.66 mg/L (group 4); 0.75 mg/L (group 5); 0.89 mg/L (group 6); 1.08 mg/L (group 7); 1.33 mg/L (group 8); 1.74 mg/L (group 9); 2.96 mg/L (group 10)	Combined Raven's Test for Rural China	Lower	Age; arsenic; iodine; lead; SES; demographics
Eswar *et al.* (2011)[w]	India	12–14	Drinking water High fluoride villages/low fluoride villages	0.29 mg/L (low); 2.45 mg/L (high)	Standard Progressive Matrices	Higher	Age; gender
Poureslami *et al.* (2011)[w]	Iran	7–9	Drinking water Endemic dental fluorosis city/reference city	0.41 mg/L (reference); 2.38 mg/L (endemic)	Persian version of Raven's Matrices Test	Higher	Gender
Shivaprakash *et al.* (2011)[w]	India	7–11	Drinking water Fluorosis severity groups (mild/moderate/severe)/all fluorosis/no fluorosis	<0.5 ppm (no fluorosis); 2.5–3.5 ppm (mild fluorosis); 2.5–3.5 ppm (moderate fluorosis); 2.5–3.5 ppm (severe fluorosis); 2.5–3.5 ppm (all fluorosis)	Raven's Colored Progressive Matrices	Higher	Health factors; SES
Xiang *et al.* (2011)	China	8–13	Serum Quartiles (Q4/Q3/Q2 and Q1)	<0.05 mg/L (Q1 and Q2 reference); 0.05–0.08 mg/L (Q3); >0.08 mg/L (Q4)	Combined Raven's Test for Rural China	Lower	Age; gender; iodine; lead; SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Seraj *et al.* (2012)[w]	Iran	6–11	Drinking water High/medium/low fluoride levels	0.8 ± 0.3 mg/L (normal); 3.1 ± 0.9 mg/L (medium); 5.2 ± 1.1 mg/L (high)	Raven's Colored Progressive Matrices	Lower	Age; gender; SES
Trivedi *et al.* (2012)[w]	India	12–13	Ground water High fluoride area/low fluoride area	0.84 ± 0.38 mg/L (low); 2.3 ± 0.87 mg/L (high)	questionnaire prepared by Professor JH Shah	Lower	Gender; SES
Nagarajappa *et al.* (2013)[w]	India	8–10	Drinking water High fluoride area/low fluoride area	0.5 mg/L (low); 2.4–3.5 mg/L (high)	Seguin Form Board Test	Higher	Gender; SES; demographics
Bai *et al.* (2014)[u*]	China	8–12	Urine Coal-burning-borne fluorosis areas (seriously-affected/lightly-affected/reference)	0.54 mg/L (reference); 0.81 mg/L (lightly-affected area); 1.96 mg/L (seriously-affected area)	Chinese Combined Raven Test-C2	Higher	SES
Karimzade *et al.* (2014)[w]	Iran	9–12	Drinking water High fluoride area/low fluoride area	0.25 mg/L (low); 3.94 mg/L (high)	Iranian version of the Raymond B Cattell test	Higher	Gender
Broadbent *et al.* (2015)[w*]	New Zealand	7–13	Drinking water Area with community water fluoridation (high)/area without community water fluoridation (low) Fluoride tablet use (ever/never) Fluoride toothpaste use (always/sometimes/never)	0.0–0.3 mg/L (low); 0.7–1.0 mg/L (high) 0 mg (never used fluoride tablets); 0.5 mg (ever used fluoride tablets) Range not specified for fluoride toothpaste use (always/sometimes/never)	Wechsler Intelligence Scale for Children-Revised	Higher	Gender; SES; low birth weight; breastfeeding
Khan *et al.* (2015)	India	6–11	Drinking water High fluoride areas (Unnao)/low fluoride areas (Tiwariganj) Fluorosis grades (normal/very mild/mild/moderate/severe)	0.19 ppm (Tiwariganj); 2.41 ppm (Unnao) Range not specified by fluorosis grades	Raven's Colored Progressive Matrices	Higher	Health factors; SES
Sebastian and Sunitha (2015)[w*]	India	10–12	Drinking water Low fluoride villages/normal fluoride villages/high fluoride villages	0.40 mg/L (low); 1.2 mg/L (normal); 2.0 mg/L (high)	Raven's Colored Progressive Matrices	Higher	Age; gender; SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Zhang *et al.* (2015b)[w*]	China	10–12	Urine, drinking water, serum High fluoride areas/reference	1.10 ± 0.67 mg/L urine; 0.63 (0.58–0.68)mg/L water; 0.06 ± 0.03 serum (reference); 2.40 ± 1.01 mg/L urine; 1.40 (1.23–1.57) mg/L water; 0.18 ± 0.11 serum (high)	Combined Raven's Test for Rural China	Lower	Age; gender; arsenic; iodine; drinking water fluoride; SES; thyroid hormone levels; COMT genotype
Das and Mondal (2016)[u]	India	6–18	Urine, drinking water Dental fluorosis (severe/moderate/mild/very mild/normal)	2.91 ± 1.76 mg/L urine; 0.069 ± 0.021 mg/kg-d drinking water (normal dental fluorosis); 2.50 ± 2.39 mg/L urine; 0.064 ± 0.004 mg/kg-d drinking water (questionable dental fluorosis); 2.58 ± 1.31 mg/L urine; 0.060 ± 0.036 mg/kg-d drinking water (very mild dental fluorosis); 2.95 ± 1.44 mg/L urine; 0.060 ± 0.030 mg/kg-d drinking water (mild dental fluorosis); 4.82 ± 3.57 mg/L urine; 0.099 ± 0.063 mg/kg-d drinking water (moderate dental fluorosis); 3.81 ± 2.51 mg/L urine; 0.093 ± 0.040 mg/kg-d drinking water (severe dental fluorosis)	Combined Raven's Test for Rural China	Higher	–
Mondal *et al.* (2016)[w]	India	10–14	Drinking water High fluoride areas/low fluoride areas	Not reported (low); 0.33–18.08 mg/L (high)	Raven Standard Theoretical Intelligence Test	Higher	SES

Table A-1. Characteristics of Studies Included in the Mean-effect Meta-analysis

Reference[1]	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Bashash *et al.* (2017) [u]	Mexico	6–12	Urine	<0.80 mg/L (reference); ≥0.80 mg/L (high)	Wechsler Abbreviated Scale of Intelligence	Lower	Age; gender; weight at birth; parity; gestational age; maternal characteristics (smoking history, marital status, age at delivery, IQ, education, cohort)
Yu *et al.* (2018)[u*,w]	China	7–13	Drinking water High fluoride/normal fluoride Urine (per 0.5-mg/L increment in each range) High/medium/low fluoride ranges	≤1 mg/L (normal); >1 mg/L (high) 0.01–1.60 mg/L (low range urinary); 1.60–2.50 mg/L (medium range urinary); 2.50–5.54 mg/L (high range urinary)	Combined Raven's Test for Rural China	Lower	Age; gender; health factors; SES
Cui *et al.* (2020)[u]	China	7–12	Urine High/medium/low fluoride levels	<1.6 mg/L (low); 1.6–2.5 mg/L (medium); >=2.5 mg/L (high)	Combined Raven's Test	Lower	Gender; arsenic; iodine

Notes:
COMT = catechol-O-methyltransferase; RoB = risk of bias; SES = socioeconomic status
[1]A "w" superscript indicates studies included in the mean effect dose-response meta-analysis using fluoride in water; a "u" superscript indicates studies included in the mean effect dose-response meta-analysis using fluoride in urine; "*" indicates studies included in the mean effect dose-response meta-analysis at levels < 1.5 mg/L.

Table A-2. Studies Excluded from Mean-effect Meta-analysis	
Reference, Country	**Reason for Exclusion**
Qin *et al.* (1990)[translated in Qin *et al.* 2008], China	Missing mean or SD of outcome measure
Yang *et al.* (1994) [translated in Yang *et al.* 2008], China	Overlapping population with Wang *et al.* (2001); Table 2 in Yang *et al.* (1994) seemed incomplete
Wang *et al.* (2005b) [translated in Wang *et al.* 2008a], China	Missing mean or SD of outcome measure
Rocha-Amador *et al.* (2007), Mexico	Missing mean or SD of outcome measure
Liu *et al.* (2000) [translated in Liu *et al.* 2008], China	Overlapping population with Lu *et al.* (2000)
Sudhir *et al.* (2009), India	Missing mean or SD of outcome measure
He and Zhang (2010), China	Missing mean or SD of outcome measure
Kang *et al.* (2011), China	Missing mean or SD of outcome measure
Saxena *et al.* (2012), India	Missing mean or SD of outcome measure
Wang *et al.* (2012), China	Overlapping population with Xiang *et al.* (2003a); used in individual-level meta-analysis
Singh *et al.* (2013), India	Missing mean or SD of outcome measure
Wei *et al.* (2014), China	Missing mean or SD of outcome measure
Choi *et al.* (2015), China	Cognitive functions other than IQ
Kundu *et al.* (2015), India	Unusual IQ scores; used only for sensitivity analysis
Aravind *et al.* (2016), India	Unusual IQ scores; used only for sensitivity analysis
Razdan *et al.* (2017), India	Unusual IQ scores; used only for sensitivity analysis
Cui *et al.* (2018), China	Missing mean or SD of outcome measure; used in individual-level meta-analysis
Green *et al.* (2019), Canada	Missing mean or SD of outcome measure; used in individual-level meta-analysis
Soto-Barreras *et al.* (2019), Mexico	Missing mean or SD of outcome measure
Zhao *et al.* (2019), China	Overlapping population with Yu *et al.* (2018), but smaller sample size
Zhou *et al.* (2019), China	Overlapping population with Yu *et al.* (2018), but smaller sample size
Till *et al.* (2020), Canada	Missing mean or SD of outcome measure; used in individual-level meta-analysis
Wang *et al.* (2020b), China	Missing mean or SD of outcome measure; used in individual-level meta-analysis
Zhao *et al.* (2020), China	Overlapping population with Yu *et al.* (2018), but smaller sample size

Notes:
SD = standard deviation

Table A-3. Characteristics of Studies Included in the Individual-level Meta-analysis

Reference	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Ding *et al.* (2011)	China	7–14	Urine	0.10–3.55 mg/L	Combined Raven's Test for Rural China	Lower	Age; arsenic; iodine; lead; SES; demographics
Zhang *et al.* (2015b)	China	10–12	Urine High fluoride area/reference	1.10 ± 0.67 mg/L (reference); 2.40 ± 1.01 mg/L (high fluoride area)	Combined Raven's Test for Rural China	Lower	Age; gender; arsenic; iodine; drinking water fluoride; SES; thyroid hormone levels; COMT genotype
Bashash *et al.* (2017)	Mexico	6–12	Urine	0.18–2.8 mg/L	Wechsler Abbreviated Scale of Intelligence	Lower	Age; gender; weight at birth; parity; gestational age; maternal characteristics (smoking history, marital status, age at delivery, IQ, education, cohort)
Cui *et al.* (2018)	China	7–12	Urine	0.8–2.0 mg/L	Combined Raven's Test for Rural China	Lower	Age; maternal education; smoking in family member; stress; anger; dopamine receptor-2 polymorphism
Yu *et al.* (2018)	China	7–13	Urine, drinking water	0.01–5.54 mg/L (urine) 0.20–3.90 mg/L (water)	Combined Raven's Test for Rural China	Lower	Age; gender; maternal education; paternal education; low birth weight
Green *et al.* (2019)	Canada	3–4	Maternal urine, maternal fluoride intake, drinking water	0.51 ± 0.36 mg/L (urine) 0.54 ± 0.44 mg/day (maternal daily fluoride intake) 0.31 ± 0.23 mg/L (water)	Wechsler Primary and Preschool Scale of Intelligence-III	Lower	Gender; city; maternal education; race/ethnicity; HOME score; prenatal secondhand smoke exposure

Table A-3. Characteristics of Studies Included in the Individual-level Meta-analysis

Reference	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric, Exposure Groups)	Range			
Till *et al.* (2020)	Canada	3–4	Residence, maternal urine, infant fluoride intake from formula, drinking water Fluoridated/nonfluoridated areas	0.64–0.70 (fluoridated), 0.38–0.42 mg/L (nonfluoridated) (urine) 0.12–0.34 (fluoridated), 0.02–0.08 mg/day (nonfluoridated) (infant formula fluoride intake) 0.58 (fluoridated), 0.13 (nonfluoridated) (water)	Wechsler Primary and Preschool Scale of Intelligence-III	Lower	Age; gender; maternal education; maternal race; HOME total score; secondhand smoke status in the child's house
Wang *et al.* (2020b)	China	7–13	Urine, drinking water	0.01–5.54 mg/L (urine) 0.20–3.90 mg/L (water)	Combined Raven's Test for Rural China	Lower	Age; gender; body mass index; maternal education; paternal education; household income; low birth weight

Notes:
COMT = catechol-O-methyltransferase; RoB = risk of bias; SES = socioeconomic status; HOME = Home Observation for Measurement of the Environment

Effects by Risk-of-bias Evaluation

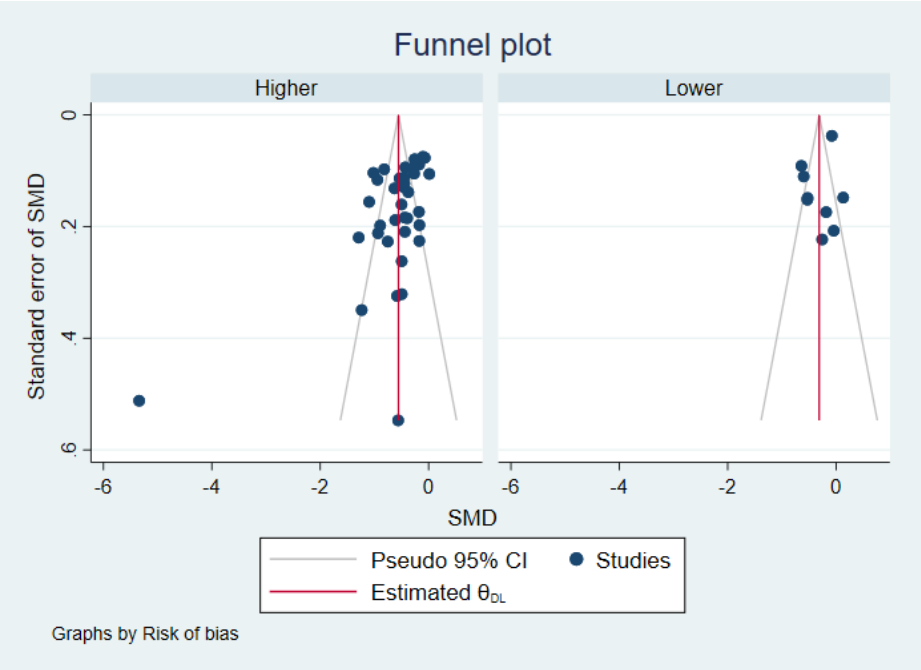


Figure A-1. Funnel Plot by Risk-of-bias Evaluation

```
Regression-based Egger test for small-study effects
Random-effects model                                  Higher RoB
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -3.63
      SE of beta1 =     0.709
                z =     -5.12
       Prob > |z| =    0.0000


Regression-based Egger test for small-study effects
Random-effects model                                  Lower RoB
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =      0.78
      SE of beta1 =     2.155
                z =      0.36
       Prob > |z| =    0.7169
```

Figure A-2. Test for Publication Bias by Risk of Bias

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     46
  Model: Random-effects                      observed =     37
 Method: DerSimonian-Laird                   imputed =      9

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.556	-0.684	-0.428
Observed + Imputed	-0.354	-0.498	-0.210

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     44
  Model: Random-effects                      observed =     37
 Method: DerSimonian-Laird                   imputed =      7

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.556	-0.684	-0.428
Observed + Imputed	-0.684	-0.831	-0.537

Figure A-3. Trim-and-fill Analysis for High Risk-of-bias Studies

Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

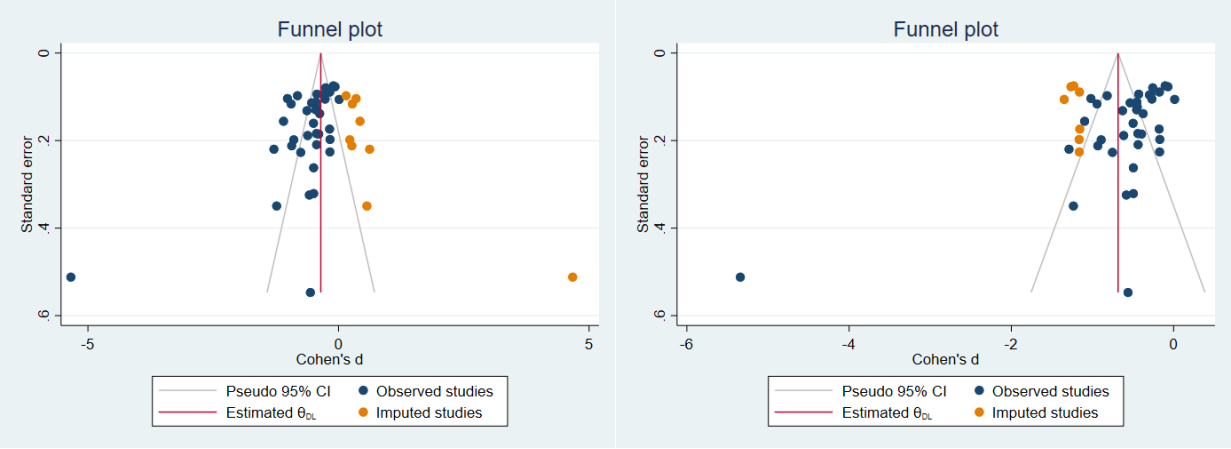


Figure A-4. Filled-in Funnel Plots for High Risk-of-bias Studies

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator (the linear estimator to the right showed no change in the pooled SMD); right panel shows random-effects pooled SMD after filling in to the left using a linear estimator (the run estimator to the left showed no change in the pooled SMD).

Effects by Gender

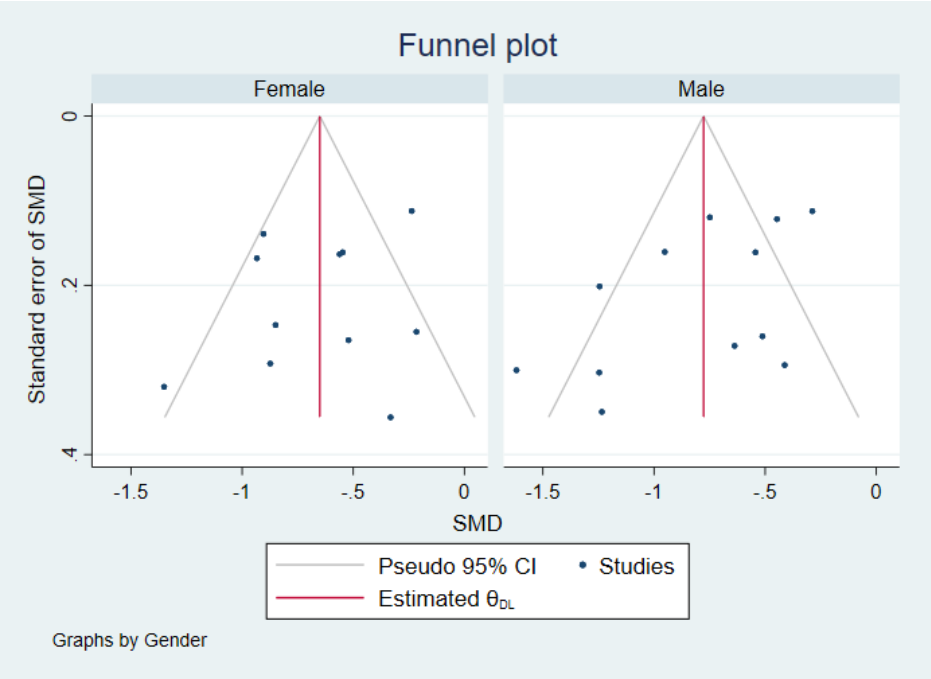


Figure A-5. Funnel Plots by Gender

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird
                                                Females
H0: beta1 = 0; no small-study effects
            beta1 =     -0.90
      SE of beta1 =     1.423
                z =     -0.63
      Prob > |z| =    0.5282

 Regression-based Egger test for small-study effects
 Random-effects model
 Method: DerSimonian-Laird
                                                Males
 H0: beta1 = 0; no small-study effects
             beta1 =     -2.60
       SE of beta1 =     1.225
                 z =     -2.12
       Prob > |z| =    0.0338
```

Figure A-6. Test for Publication Bias by Gender

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration                          Number of studies =     14
  Model: Random-effects                     observed =     12
 Method: DerSimonian-Laird                  imputed =      2

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.777	-0.994	-0.560
Observed + Imputed	-0.681	-0.900	-0.461

Figure A-7. Trim-and-fill Analysis for Studies in Boys Using Linear and Run Estimators

Filling in to the right using a run estimator or to the left using a linear or a run estimator showed no change in the pooled SMD.

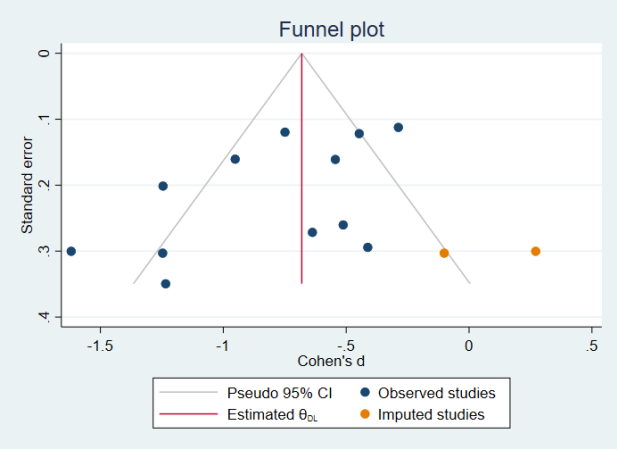


Figure A-8. Filled-in Funnel Plots for Studies in Boys

Panel shows funnel plot filled in to the right using a linear estimator. Filling in to the right using a run estimator or to the left using a linear or a run estimator showed no change in the pooled SMD.

Effects by Age Group

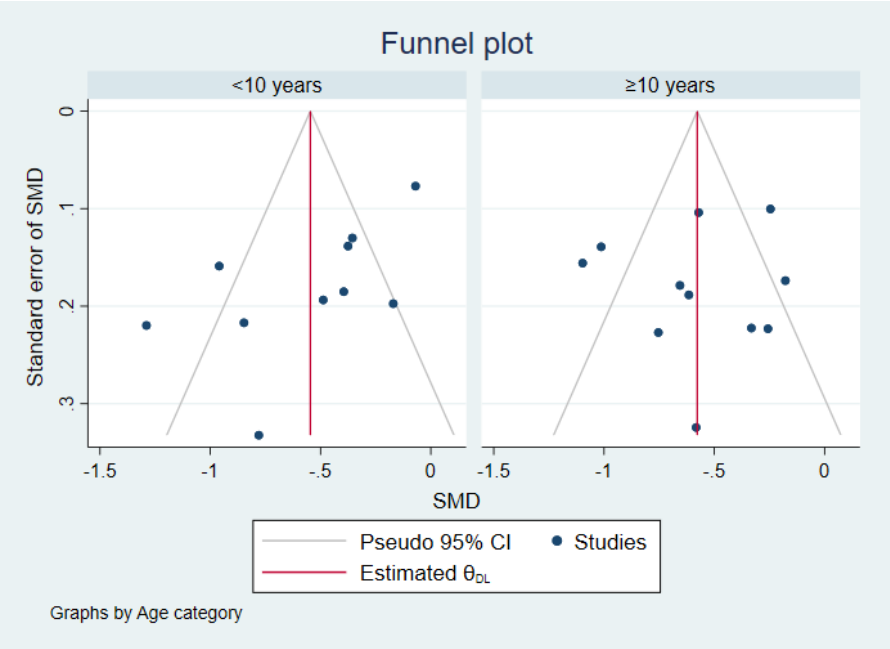


Figure A-9. Funnel Plot by Age Group

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird                          < 10 years

H0: beta1 = 0; no small-study effects
            beta1 =     -3.53
      SE of beta1 =     1.631
                z =     -2.16
       Prob > |z| =    0.0307


Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird                          ≥ 10 years

H0: beta1 = 0; no small-study effects
            beta1 =      0.32
      SE of beta1 =     1.883
                z =      0.17
       Prob > |z| =    0.8638
```

Figure A-10. Test for Publication Bias by Age Group

Note: Although suggestive of publication bias in the less-than-10 age group, filling in to the right or left using a linear or run estimator showed no change in the pooled SMD.

Effects by Country

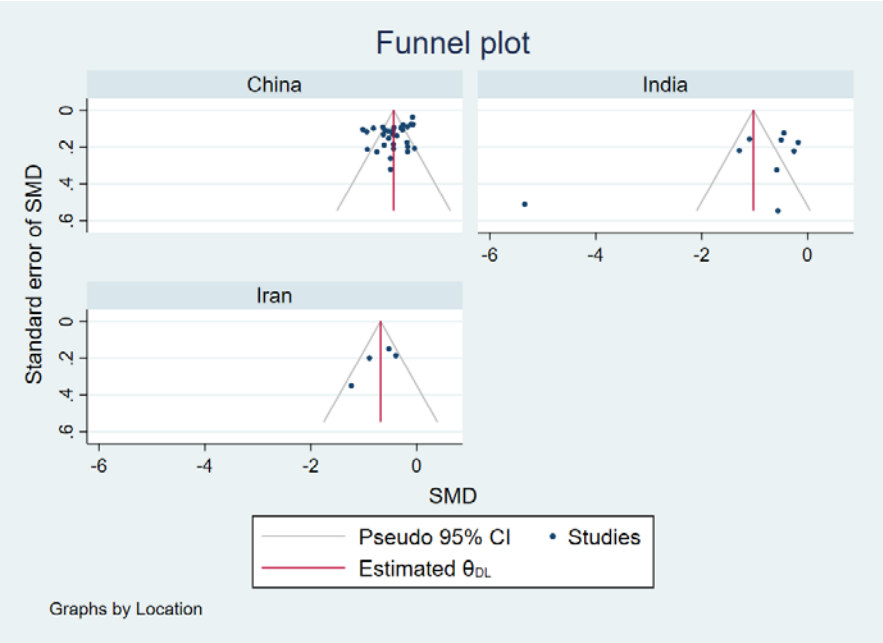


Figure A-11. Funnel Plot by Country

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird
                                        China

H0: beta1 = 0; no small-study effects
            beta1 =     -0.83
      SE of beta1 =     0.875
                z =     -0.95
       Prob > |z| =    0.3399
                                        India

H0: beta1 = 0; no small-study effects
            beta1 =     -5.64
      SE of beta1 =     1.770
                z =     -3.18
       Prob > |z| =    0.0015
                                        Iran

H0: beta1 = 0; no small-study effects
            beta1 =     -3.78
      SE of beta1 =     2.062
                z =     -1.83
       Prob > |z| =    0.0669
```

Figure A-12. Test for Publication Bias by Country

Nonparametric trim-and-fill analysis of publication bias **India**
Linear estimator, imputing on the left

Iteration — Number of studies = 12
Model: Random-effects — observed = 9
Method: DerSimonian-Laird — imputed = 3

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-1.023	-1.540	-0.506
Observed + Imputed	-1.491	-2.154	-0.828

Figure A-13. Trim-and-fill Analysis for Studies in India

Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

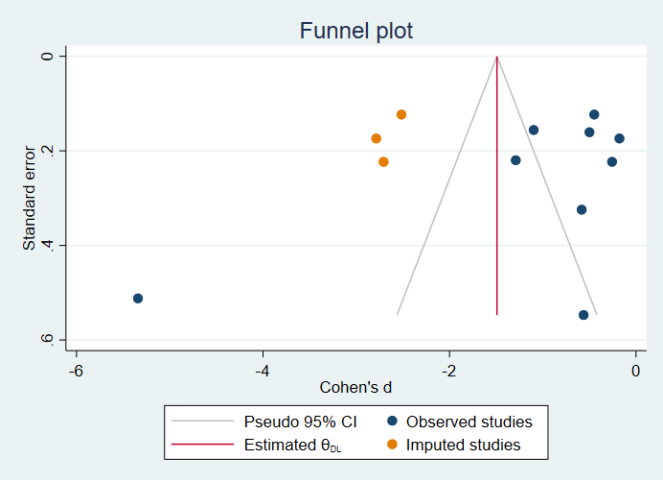


Figure A-14. Filled-in Funnel Plot for Studies in India

Panel shows funnel plot filled in to the left using a linear estimator. Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

Effect by Assessment Type

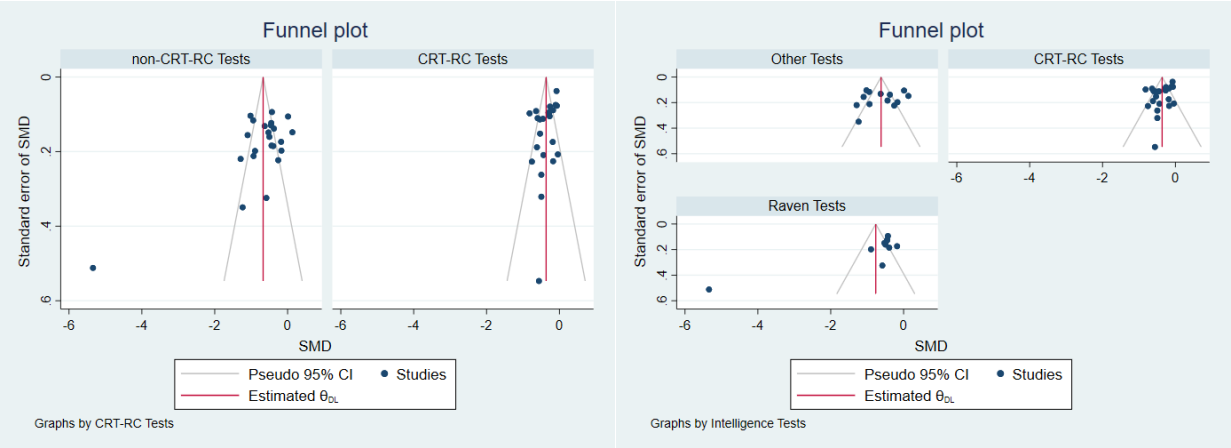


Figure A-15. Funnel Plot by CRT-RC-type Test

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird
                                        CRT-CR Tests
H0: beta1 = 0; no small-study effects
            beta1 =     -0.93
      SE of beta1 =     0.707
                z =     -1.31
       Prob > |z| =    0.1904
                                    non CRT-RC Tests
H0: beta1 = 0; no small-study effects
            beta1 =     -6.91
      SE of beta1 =     1.277
                z =     -5.41
       Prob > |z| =    0.0000
                                    Raven's Tests
H0: beta1 = 0; no small-study effects
            beta1 =     -9.17
      SE of beta1 =     1.488
                z =     -6.16
       Prob > |z| =    0.0000
                                    Other Tests
H0: beta1 = 0; no small-study effects
            beta1 =     -2.28
      SE of beta1 =     2.322
                z =     -0.98
       Prob > |z| =    0.3259
```

Figure A-16. Test for Publication Bias by Assessment Type

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     30
  Model: Random-effects                      observed =     23
 Method: DerSimonian-Laird                   imputed =       7

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.672	-0.874	-0.469
Observed + Imputed	-0.920	-1.152	-0.687

Figure A-17. Trim-and-fill Analysis in Non-CRT-RC Tests

Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

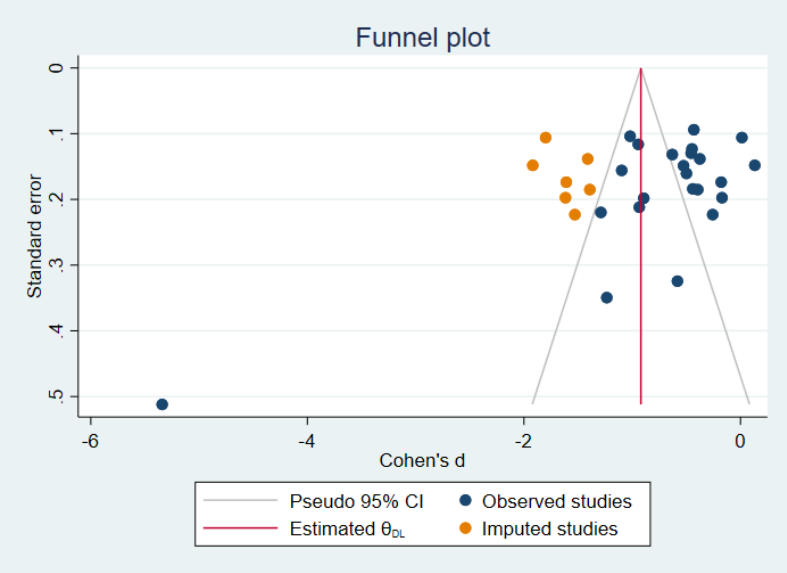


Figure A-18. Filled-in Funnel Plot to Eliminate Publication Bias in Non-CRT-RC Tests

Panel shows funnel plot filled in to the left using a linear estimator. Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     14
  Model: Random-effects                      observed =     10
 Method: DerSimonian-Laird                   imputed =      4

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.763	-1.098	-0.428
Observed + Imputed	-1.215	-1.676	-0.754

Figure A-19. Trim-and-fill Analysis for Raven-type Tests

Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

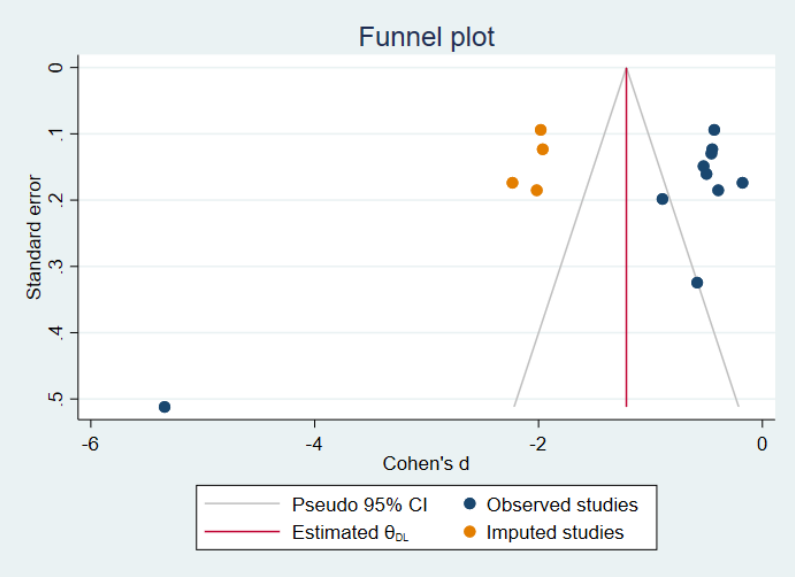


Figure A-20. Filled-in Funnel Plot to Eliminate Publication Bias for Raven-type Tests

Panel shows funnel plot filled in to the left using a linear estimator. Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

Effect by Exposure Type

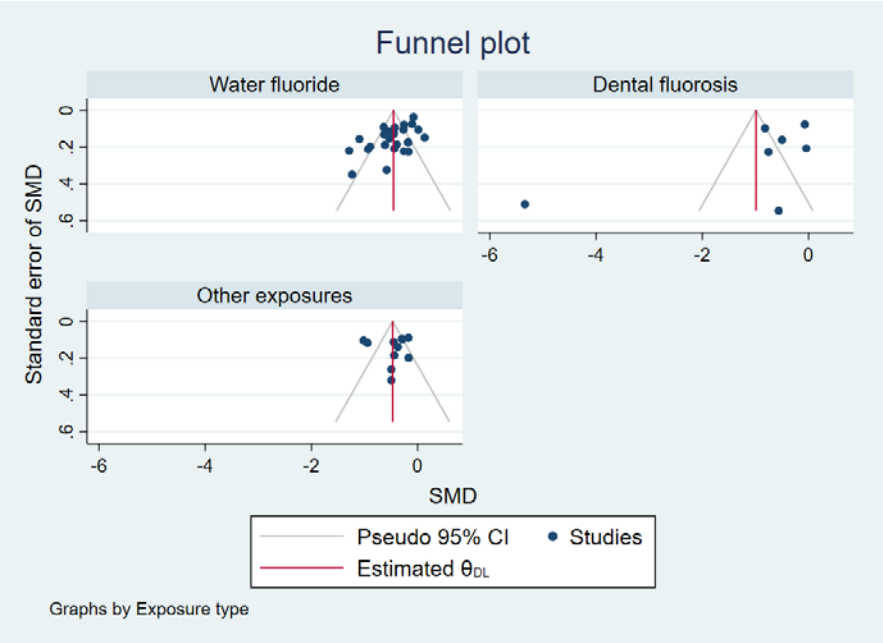


Figure A-21. Funnel Plot by Exposure Type

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird
                                        Water fluoride

H0: beta1 = 0; no small-study effects
            beta1 =     -2.44
      SE of beta1 =     0.837
                z =     -2.92
       Prob > |z| =    0.0035
                                        Dental fluorosis
H0: beta1 = 0; no small-study effects
            beta1 =     -5.81
      SE of beta1 =     1.649
                z =     -3.52
       Prob > |z| =    0.0004

                                        Other exposures
H0: beta1 = 0; no small-study effects
            beta1 =      0.21
      SE of beta1 =     1.609
                z =      0.13
       Prob > |z| =    0.8954
```

Figure A-22. Test for Publication Bias by Exposure Type

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Water fluoride

Iteration	Number of studies = 30
Model: Random-effects	observed = 28
Method: DerSimonian-Laird	imputed = 2

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.455	-0.573	-0.336
Observed + Imputed	-0.415	-0.533	-0.296

Dental fluorosis

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration	Number of studies = 9
Model: Random-effects	observed = 7
Method: DerSimonian-Laird	imputed = 2

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.989	-1.566	-0.412
Observed + Imputed	-1.510	-2.467	-0.553

Figure A-23. Trim-and-fill Analysis for Water Fluoride and Dental Fluorosis Exposures

For water fluoride, filling in to the right using a linear estimator or to the left using a linear or a run estimator showed no change in the pooled SMD. For dental fluorosis, filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

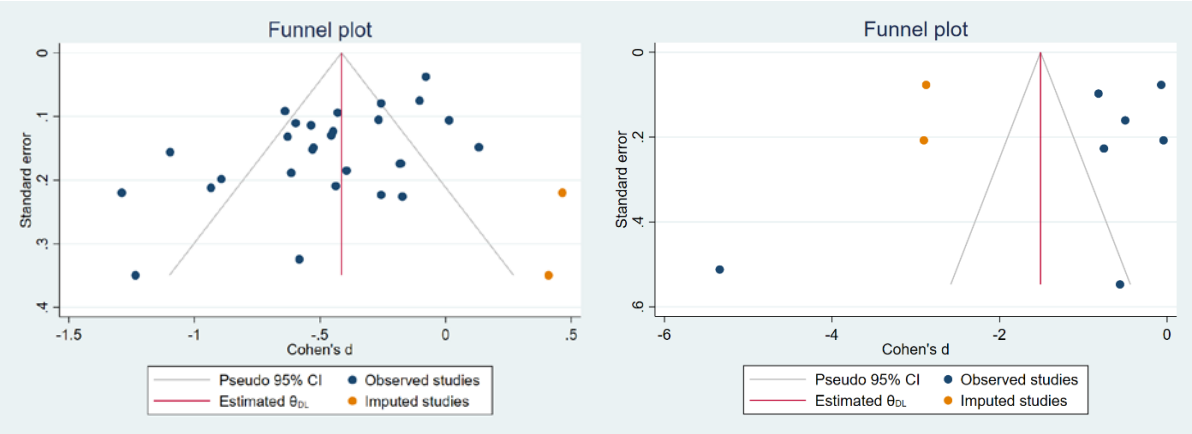


Figure A-24. Filled-in Funnel Plots to Eliminate Publication Bias for Water Fluoride (Left Panel) and Dental Fluorosis (Right Panel) Studies

For water fluoride, panel shows funnel plot filled in to the right using a run estimator. Filling in to the right using a linear estimator or to the left using a linear or a run estimator showed no change in the pooled SMD. For dental fluorosis, panel shows funnel plot filled in to the left using a linear estimator. Filling in to the right using a linear or a run estimator or to the left using a run estimator showed no change in the pooled SMD.

Sensitivity Analysis: Any Exposure Group Compared to Reference Group

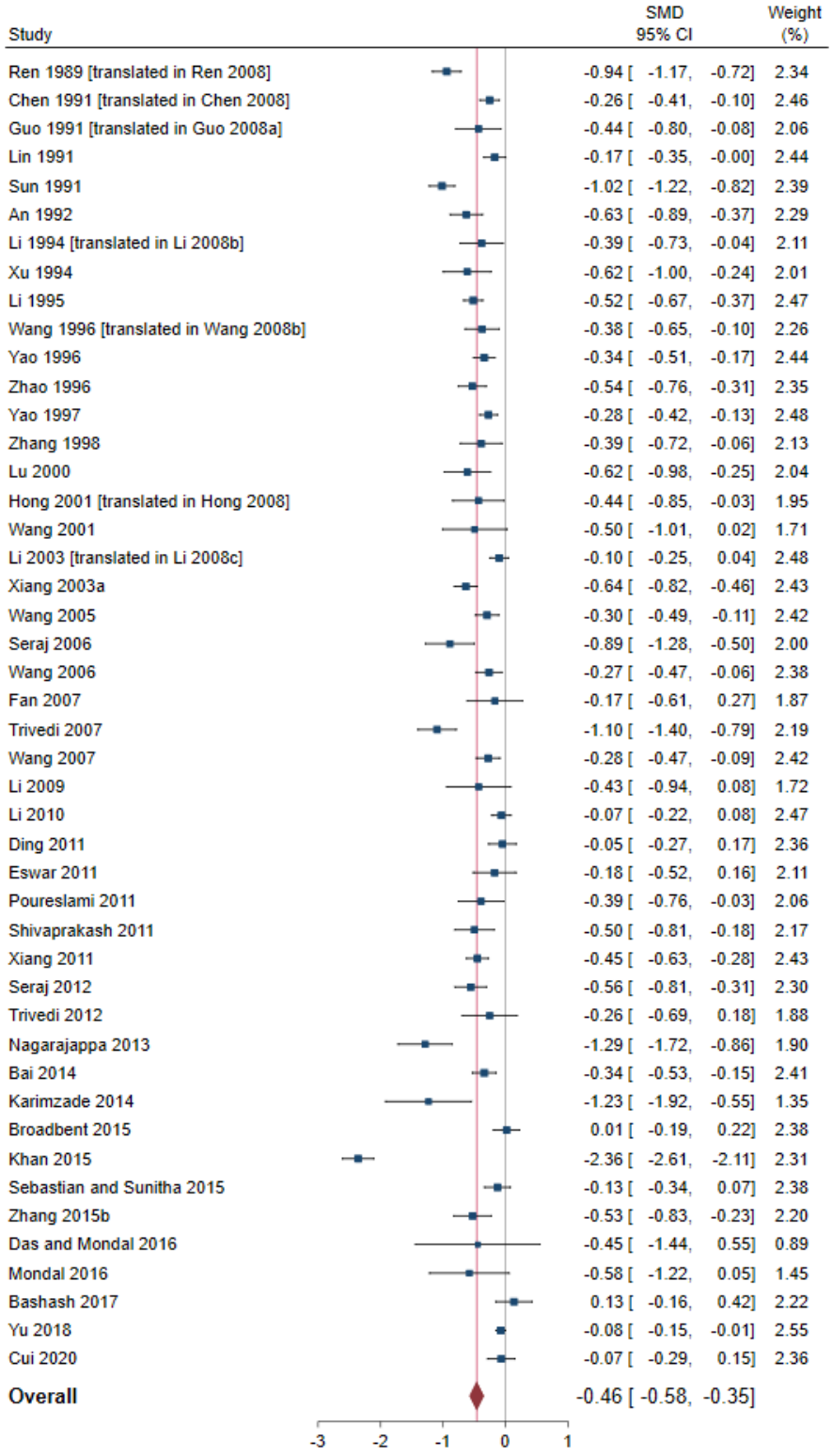


Figure A-25. Association Between Fluoride Exposure and IQ Scores in Children Using Any Exposure Versus Reference

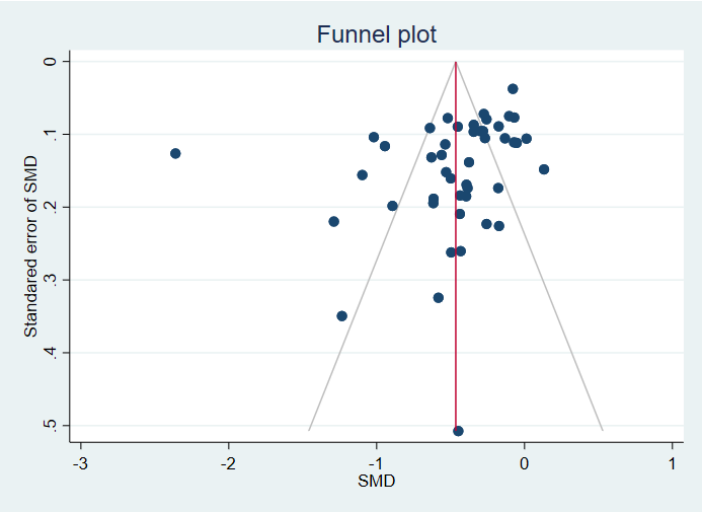


Figure A-26. Funnel Plot in Sensitivity Analysis Using Any Exposure Versus Reference

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.42
      SE of beta1 =     0.805
                z =     -1.76
       Prob > |z| =    0.0785

Begg's test for small-study effects

Kendall's score =   -237.00
    SE of score =   105.617
                z =     -2.25
     Prob > |z| =    0.0255
```

Figure A-27. Test for Publication Bias in Sensitivity Analysis Using Any Exposure Versus Reference

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     52
  Model: Random-effects                      observed =     46
 Method: DerSimonian-Laird                   imputed =      6

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.465	-0.580	-0.349
Observed + Imputed	-0.340	-0.475	-0.206

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     58
  Model: Random-effects                      observed =     46
 Method: DerSimonian-Laird                   imputed =      12

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Cohen's d	[95% Conf.	Interval]
Observed	-0.465	-0.580	-0.349
Observed + Imputed	-0.608	-0.738	-0.479

Figure A-28. Trim-and-fill Analysis in Sensitivity Analysis Using Any Exposure Versus Reference

Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

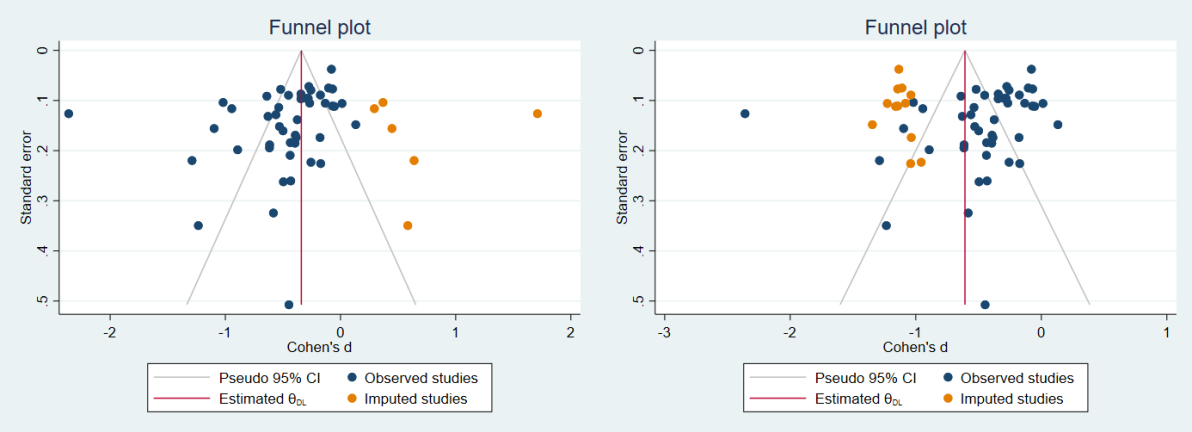


Figure A-29. Filled-in Funnel Plots to Eliminate Publication Bias in Sensitivity Analysis Using Any Exposure Versus Reference

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator (the linear estimator to the right showed no change in the pooled SMD); right panel shows random-effects pooled SMD after filling in to the left using a linear estimator (the run estimator to the left showed no change in the pooled SMD).

Sensitivity Analysis: Any Exposure Group Compared to Reference Group Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)

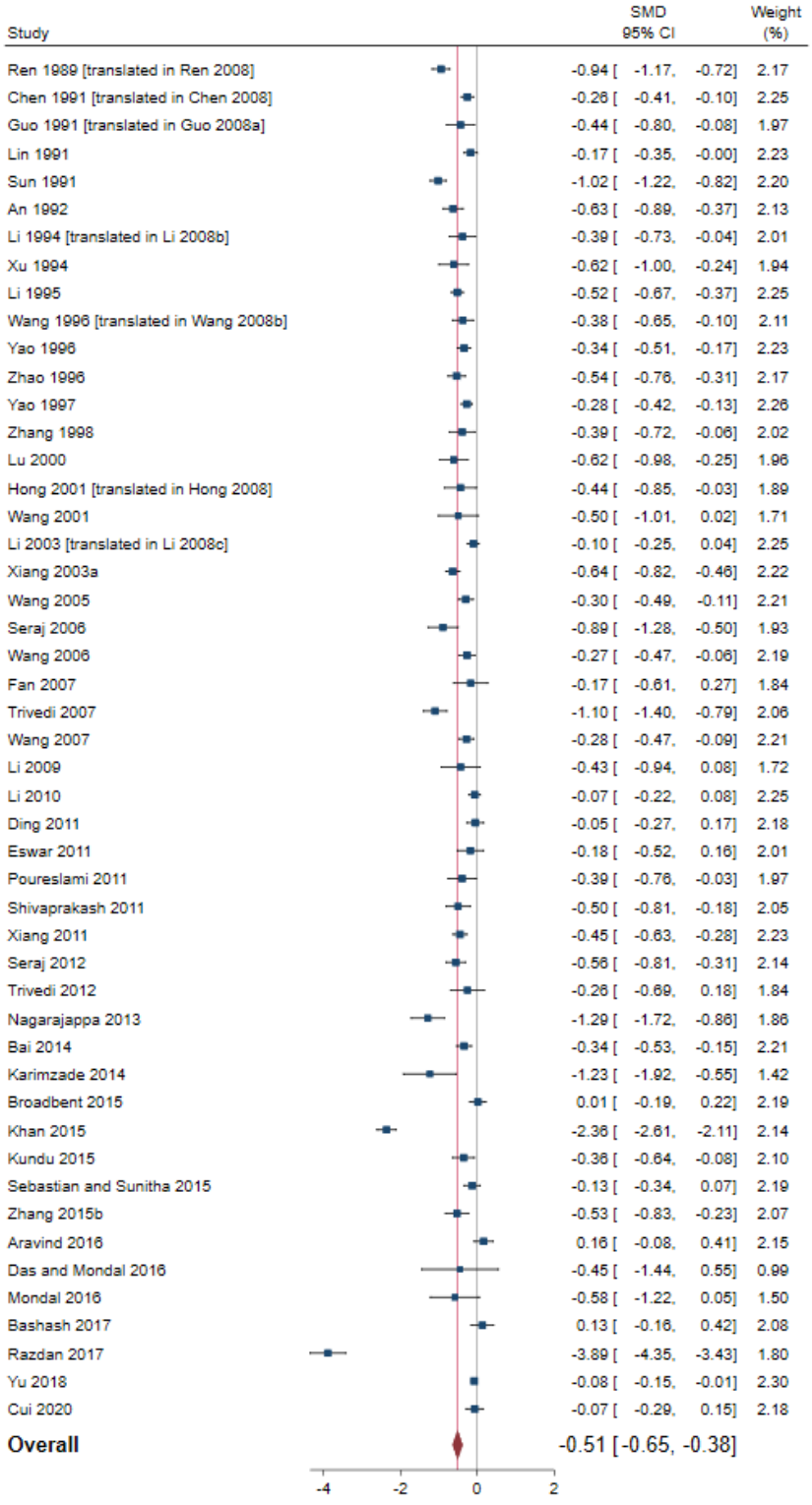


Figure A-30. Association Between Fluoride Exposure and IQ Scores in Children in Analysis Using Any Exposure Group Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

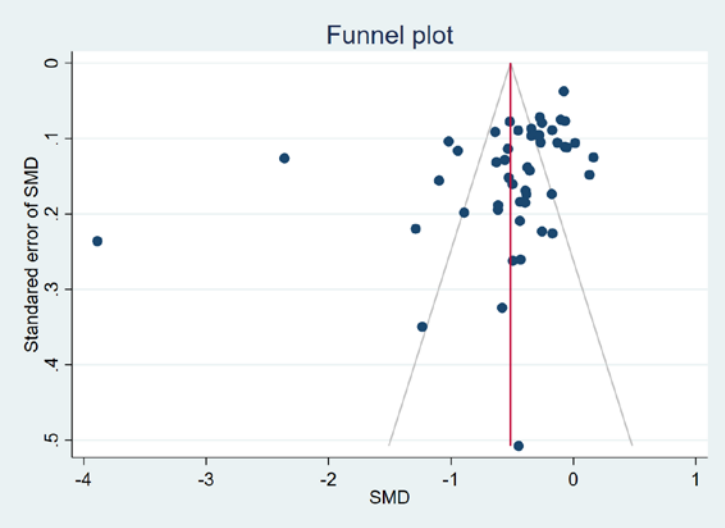


Figure A-31. Funnel Plot in Sensitivity Analysis Using Any Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -2.46
      SE of beta1 =     0.890
                z =     -2.77
      Prob > |z| =    0.0056
```

Figure A-32. Test for Publication Bias in Sensitivity Analysis Using Any Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     56
  Model: Random-effects                      observed =     49
 Method: DerSimonian-Laird                   imputed =      7

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |   Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.514      -0.645      -0.383
  Observed + Imputed |     -0.329      -0.485      -0.173
------------------------------------------------------------------------------
```

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     65
  Model: Random-effects                      observed =     49
 Method: DerSimonian-Laird                   imputed =      16

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |   Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.514      -0.645      -0.383
  Observed + Imputed |     -0.730      -0.880      -0.579
------------------------------------------------------------------------------
```

Figure A-33. Trim-and-fill Analysis in Sensitivity Analysis Using Any Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

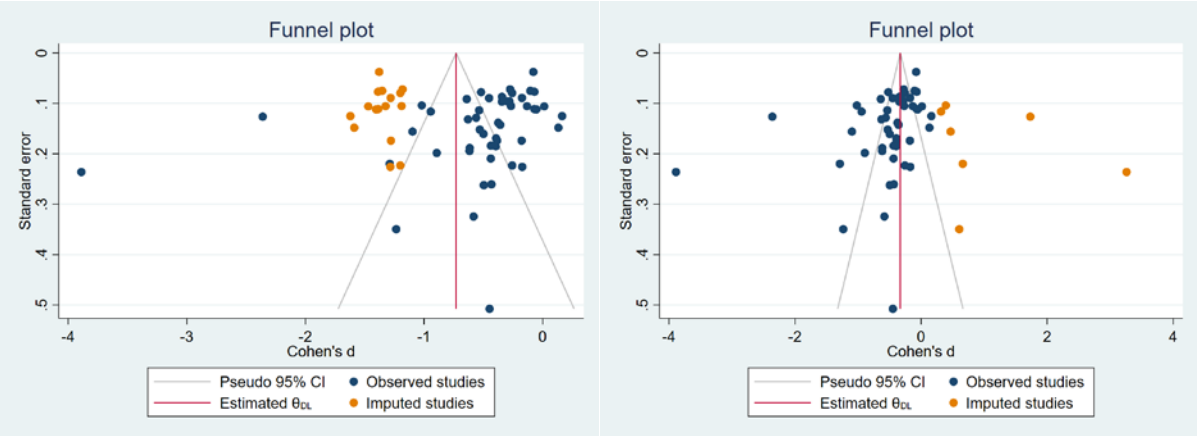


Figure A-34. Filled-in Funnel Plots to Eliminate Publication Bias in Sensitivity Analysis Using Any Exposure Versus Reference

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator; right panel shows random-effects pooled SMD after filling in to the left using a linear estimator. Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

Sensitivity Analysis: Highest Exposure Group Compared to Reference Group Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)

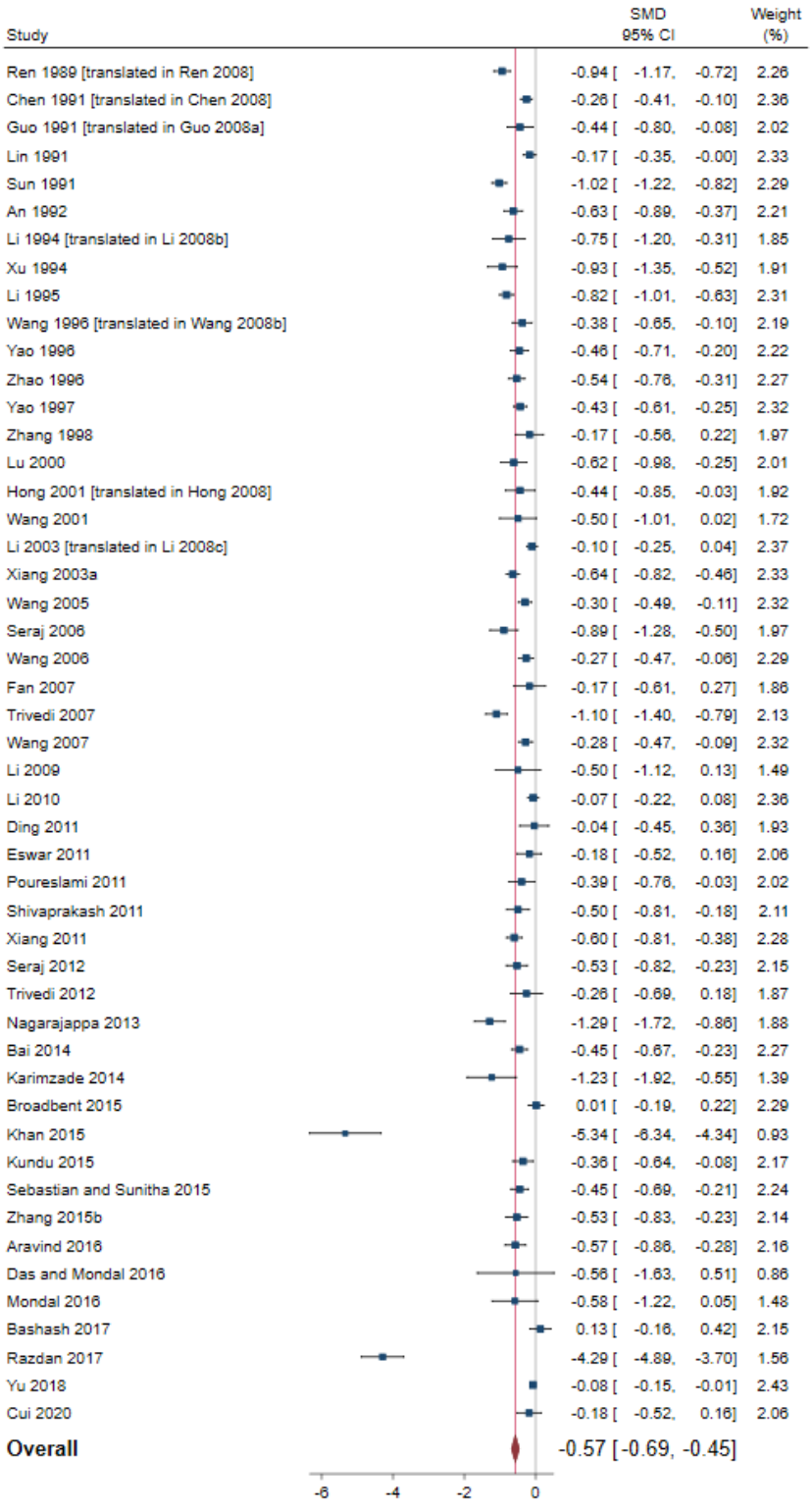


Figure A-35. Association Between Fluoride Exposure and IQ Scores in Children in Analysis Using Highest Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

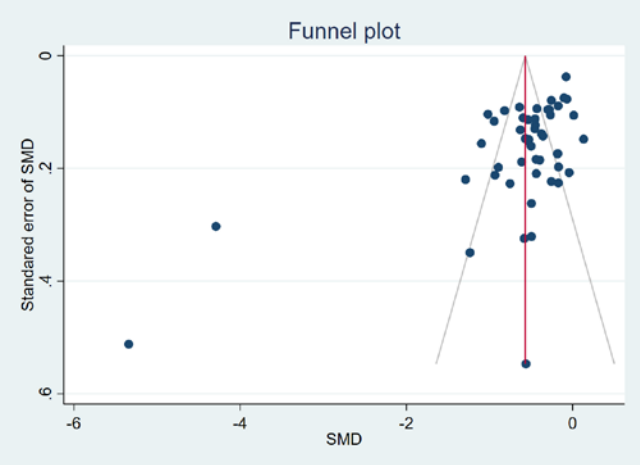


Figure A-36. Funnel Plot in Sensitivity Analysis Using Highest Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -4.47
      SE of beta1 =     0.696
                z =     -6.43
      Prob > |z| =    0.0000
```

Figure A-37. Test for Publication Bias in Sensitivity Analysis Using Highest Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     54
  Model: Random-effects                      observed =     49
 Method: DerSimonian-Laird                   imputed =      5

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |  Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.570      -0.693      -0.447
  Observed + Imputed |     -0.422      -0.563      -0.281
------------------------------------------------------------------------------
```

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     66
  Model: Random-effects                      observed =     49
 Method: DerSimonian-Laird                   imputed =      17

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |  Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.570      -0.693      -0.447
  Observed + Imputed |     -0.838      -0.999      -0.676
------------------------------------------------------------------------------
```

Figure A-38. Trim-and-fill Analysis in Sensitivity Analysis Using Highest Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

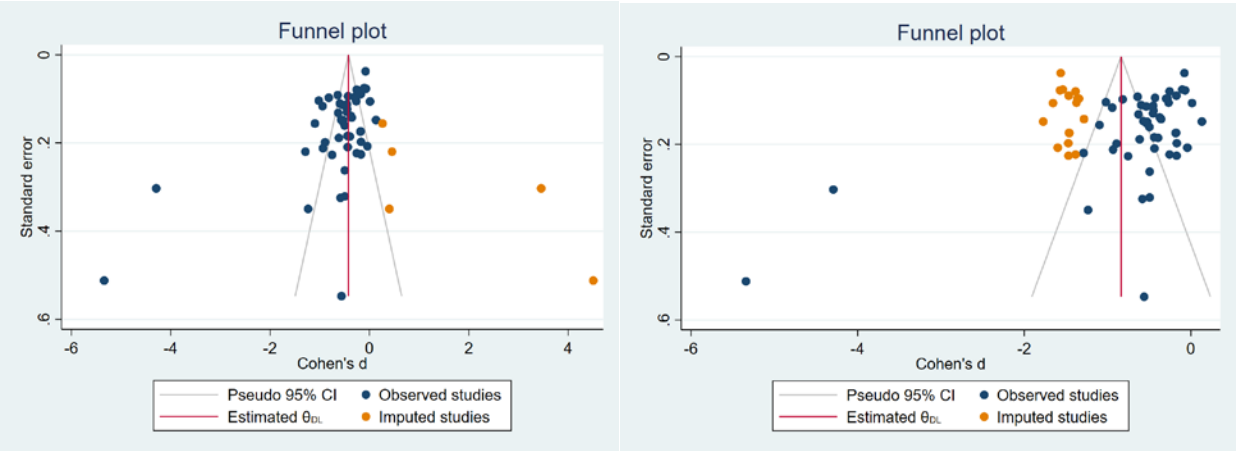


Figure A-39. Filled-in Funnel Plots to Eliminate Publication Bias in Sensitivity Analysis Using Highest Exposure Versus Reference [Including Aravind *et al.* (2016), Kundu *et al.* (2015), and Razdan *et al.* (2017)]

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator; right panel shows random-effects pooled SMD after filling in to the left using a linear estimator. Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

Sensitivity Analysis: Excluding Lin *et al.* (1991)[1]

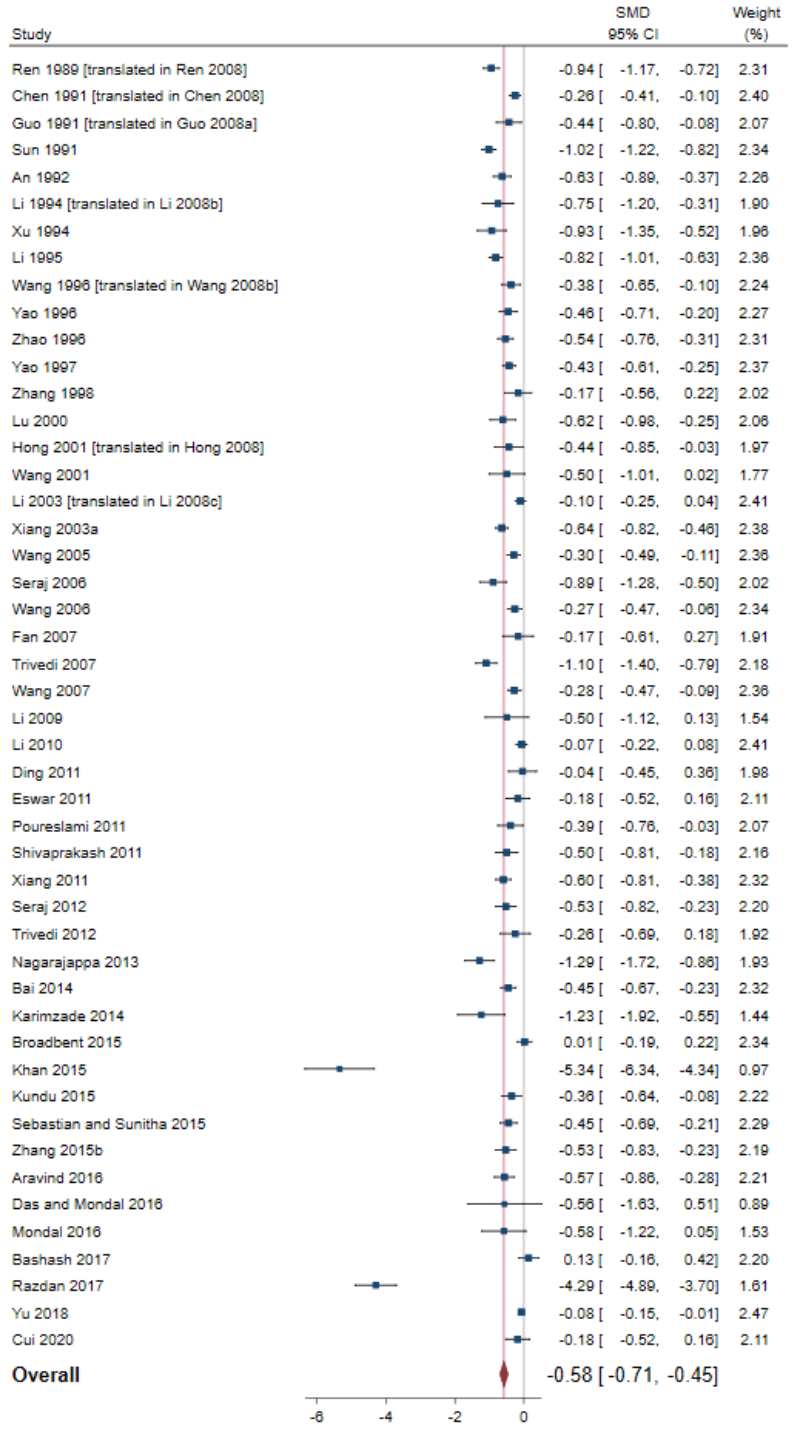


Figure A-40. Association Between Fluoride Exposure and IQ Scores in Children in Analysis Using Highest Exposure Versus Reference [Excluding Lin *et al.* (1991)]

[1]Lin *et al.* (1991): ICF calculated standard errors based on p-values.

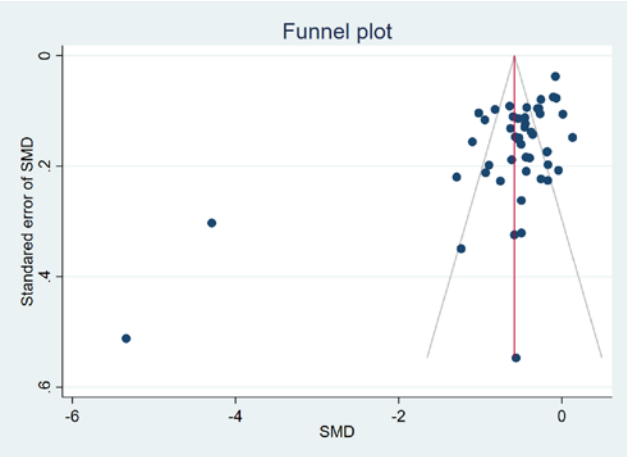


Figure A-41. Funnel Plot in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Lin *et al.* (1991)]

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -4.47
      SE of beta1 =     0.711
                z =     -6.29
      Prob > |z| =    0.0000
```

Figure A-42. Test for Publication Bias in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Lin *et al.* (1991)]

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     53
  Model: Random-effects                     observed =     48
 Method: DerSimonian-Laird                  imputed =      5

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |   Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.581      -0.707      -0.454
  Observed + Imputed |     -0.429      -0.573      -0.284
------------------------------------------------------------------------------
```

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the left

Iteration                           Number of studies =     65
  Model: Random-effects                     observed =     48
 Method: DerSimonian-Laird                  imputed =      17

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |   Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |     -0.581      -0.707      -0.454
  Observed + Imputed |     -0.855      -1.021      -0.689
------------------------------------------------------------------------------
```

Figure A-43. Trim-and-fill Analysis in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Lin *et al.* (1991)]

Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

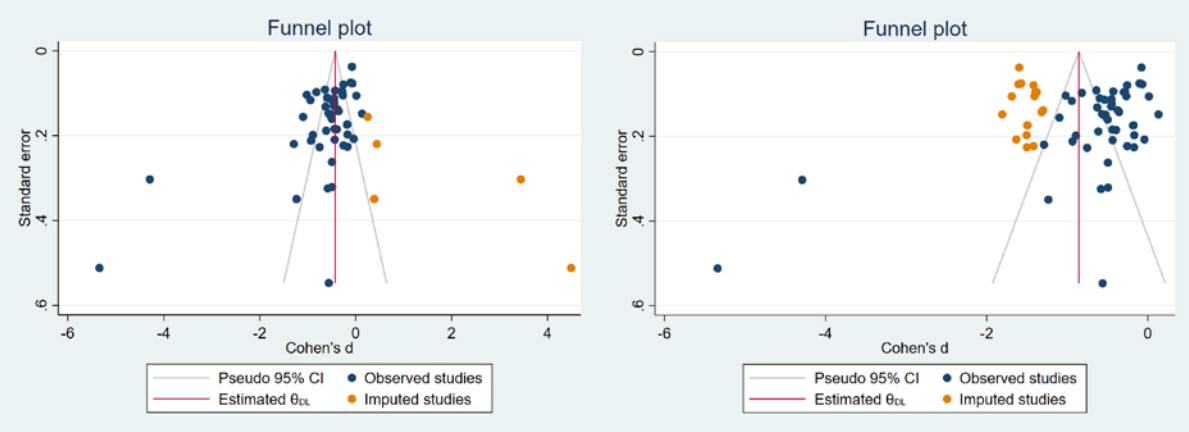


Figure A-44. Filled-in Funnel Plots to Eliminate Publication Bias in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Lin *et al.* (1991)]

Left panel shows the random-effects pooled SMD after filling in to the right using a run estimator; right panel shows random-effects pooled SMD after filling in to the left using a linear estimator. Filling in to the right using a linear estimator or to the left using a run estimator showed no change in the pooled SMD.

Sensitivity Analysis: Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)

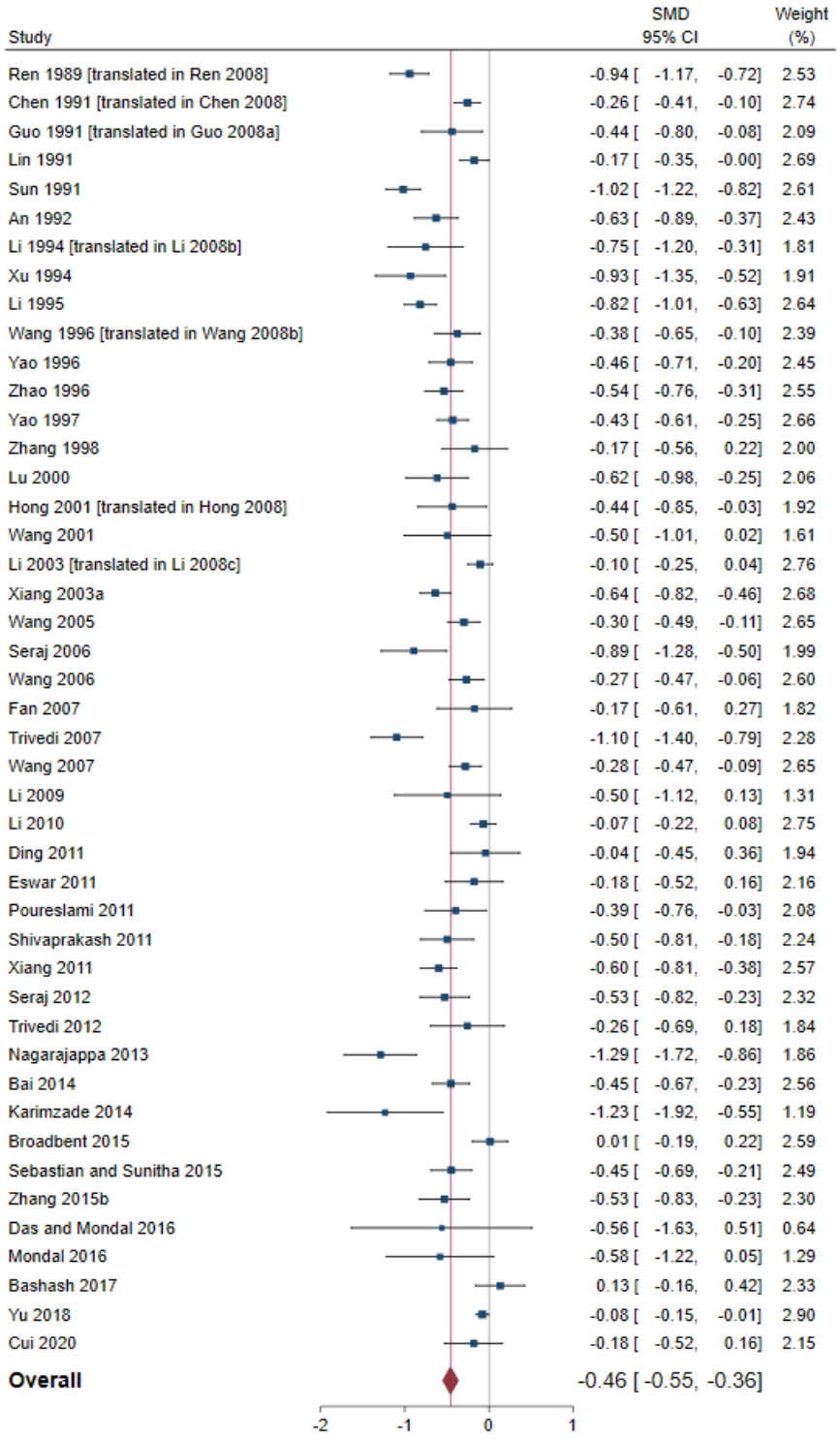


Figure A-45. Association Between Fluoride Exposure and IQ Scores in Children in Analysis Using Highest Exposure Versus Reference [Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)]

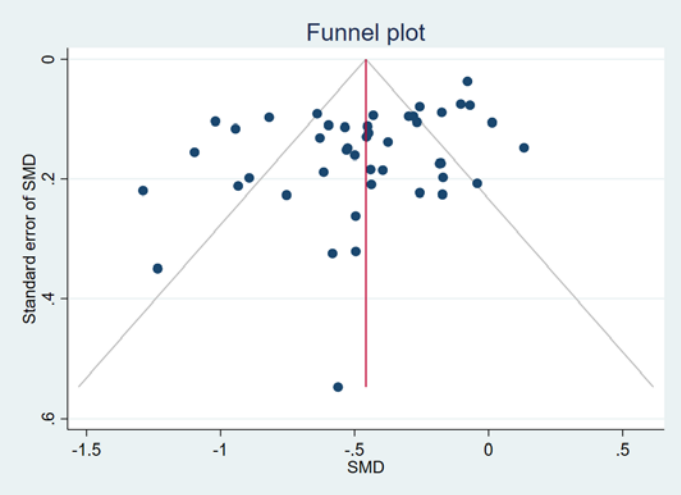


Figure A-46. Funnel Plot in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)]

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.30
      SE of beta1 =     0.659
                z =     -1.98
      Prob > |z| =    0.0481
```

Figure A-47. Test for Publication Bias in Sensitivity Analysis Using Any Exposure Versus Reference [Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)]

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                           Number of studies =     48
  Model: Random-effects                      observed =     45
 Method: DerSimonian-Laird                   imputed =       3

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird

------------------------------------------------------------------------------
             Studies |        Cohen's d    [95% Conf. Interval]
---------------------+--------------------------------------------------------
            Observed |           -0.458      -0.554      -0.361
  Observed + Imputed |           -0.417      -0.514      -0.319
------------------------------------------------------------------------------
```

Figure A-48. Trim-and-fill Analysis in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)]

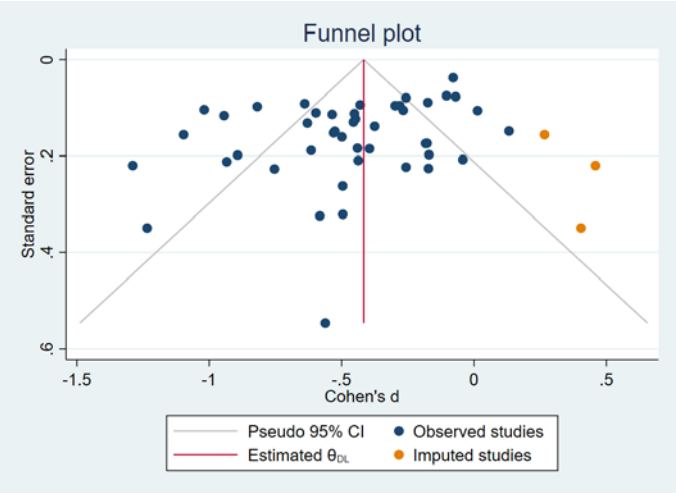


Figure A-49. Filled-in Funnel Plot to Eliminate Publication Bias in Sensitivity Analysis Using Highest Exposure Versus Reference [Excluding Aravind *et al.* (2016), Kundu *et al.* (2015), Razdan *et al.* (2017), and Khan *et al.* (2015)]

Panel shows the random-effects pooled SMD after filling in to the right using a run estimator. Filling in to the right using a linear estimator or to the left using a linear or a run estimator showed no change in the pooled SMD.

Attachment B. Subgroup and Sensitivity Analyses (Aim 2)

Table B-1. Characteristics of Studies Included in the Individual-level Meta-analysis

Reference	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric)	Range			
Ding *et al.* (2011)	China	7–14	Urine	0.10−3.55 mg/L	Combined Raven's Test for Rural China	Lower	Age; arsenic; iodine; lead; SES; demographics
Zhang *et al.* (2015b)	China	10–12	Urine	1.10 ± 0.67 mg/L (reference); 2.40 ± 1.01 mg/L (high fluoride area)	Combined Raven's Test for Rural China	Lower	Age; gender; arsenic; iodine; drinking water fluoride; SES; thyroid hormone levels; COMT genotype
Bashash *et al.* (2017)	Mexico	6–12	Urine	0.18−2.8 mg/L	Wechsler Abbreviated Scale of Intelligence	Lower	Age; gender; weight at birth; parity; gestational age; maternal characteristics (smoking history, marital status, age at delivery, IQ, education, and cohort)
Cui *et al.* (2018)	China	7–12	Urine	0.8−2.0 mg/L	Combined Raven's Test for Rural China	Lower	Age; maternal education; smoking in family member; stress; anger; dopamine receptor-2 polymorphism
Yu *et al.* (2018)	China	7–13	Urine, drinking Water	0.01−5.54 mg/L (urine) 0.20−3.90 mg/L (water)	Combined Raven's Test for Rural China	Lower	Age; gender; maternal education; paternal education; low birth weight

Table B-1. Characteristics of Studies Included in the Individual-level Meta-analysis

Reference	Study Location	Age Range (Years)	Fluoride Exposure		Intelligence Assessment	Overall RoB Rating	Confounders Considered
			Assessment (Metric)	Range			
Green *et al.* (2019)	Canada	3–4	Maternal urine, maternal fluoride intake, drinking water	0.51 ± 0.36 mg/L (urine) 0.54 ± 0.44 mg/day (maternal daily fluoride intake) 0.31 ± 0.23 mg/L (water)	Wechsler Primary and Preschool Scale of Intelligence-III	Lower	Gender; city; maternal education; race/ethnicity; HOME score; prenatal secondhand smoke exposure
Till *et al.* (2020)	Canada	3–4	Residence (fluoridated/nonfluoridated cities), maternal urine, infant fluoride intake from formula, drinking water	0.64−0.70 mg/L (fluoridated), 0.38−0.42 mg/L (nonfluoridated) (urine) 0.12−0.34 mg/day (fluoridated), 0.02−0.08 mg/day (nonfluoridated) (infant formula fluoride intake) 0.58 mg/L (fluoridated), 0.13 mg/L (nonfluoridated) (water)	Wechsler Primary and Preschool Scale of Intelligence-III	Lower	Age; gender; maternal education; maternal race; HOME total score; second-hand smoke status in the child's house
Wang *et al.* (2020b)	China	7–13	Urine, drinking water	0.01−5.54 mg/L (urine) 0.20−3.90 mg/L (water)	Combined Raven's Test for Rural China	Lower	Age; gender; body mass index; maternal education; paternal education; household income; low birth weight

Sensitivity Analysis for Individual-level Studies: No Pooling for Yu *et al.* (2018)

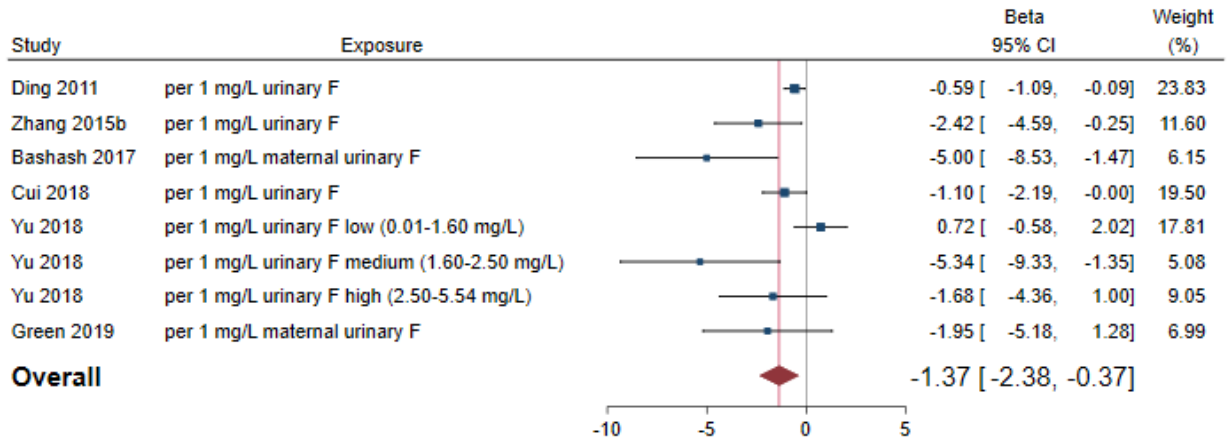


Figure B-1. Association Between Individual-level Fluoride Exposure and IQ Scores in Children [No Pooling for Yu *et al.* (2018)]

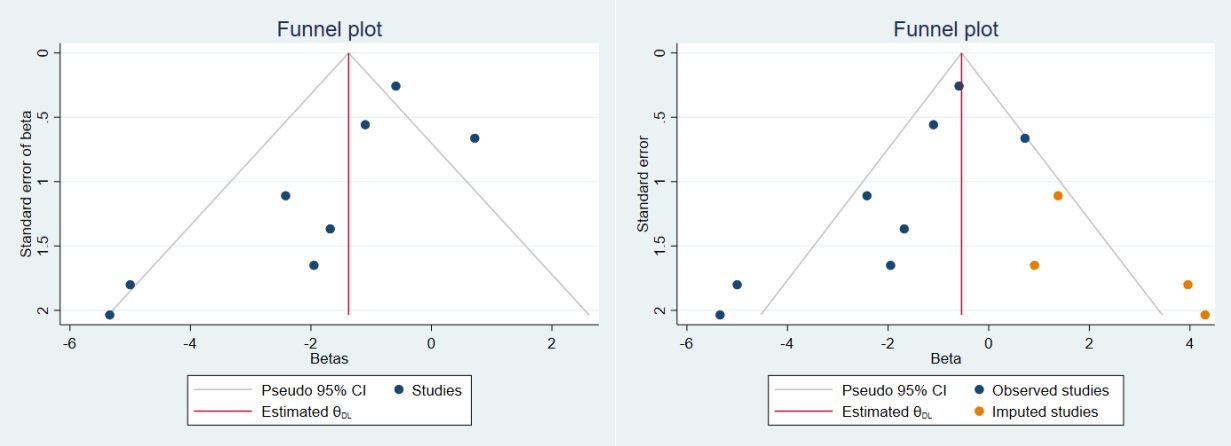


Figure B-2. Funnel Plots of Studies with Individual-level Exposures and Filled-in Funnel Plot to Eliminate Publication Bias [No Pooling for Yu *et al.* (2018)]

Right panel shows funnel plot filled in to the right using a linear estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -2.21
      SE of beta1 =     0.801
                z =     -2.75
      Prob > |z| =    0.0059
```

Figure B-3. Test for Publication Bias for Studies with Individual-level Exposures

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration
Model: Random-effects
Method: DerSimonian-Laird

Number of studies = 12
observed = 8
imputed = 4

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.374	-2.379	-0.369
Observed + Imputed	-0.540	-1.567	0.486

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration
Model: Random-effects
Method: DerSimonian-Laird

Number of studies = 9
observed = 8
imputed = 1

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.374	-2.379	-0.369
Observed + Imputed	-1.188	-2.220	-0.155

Figure B-4. Trim-and-fill Analysis for Studies with Individual-level Exposures

Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

Sensitivity Analysis for Individual-level Studies: Effect by Exposure Type, No Pooling for Yu *et al.* (2018)

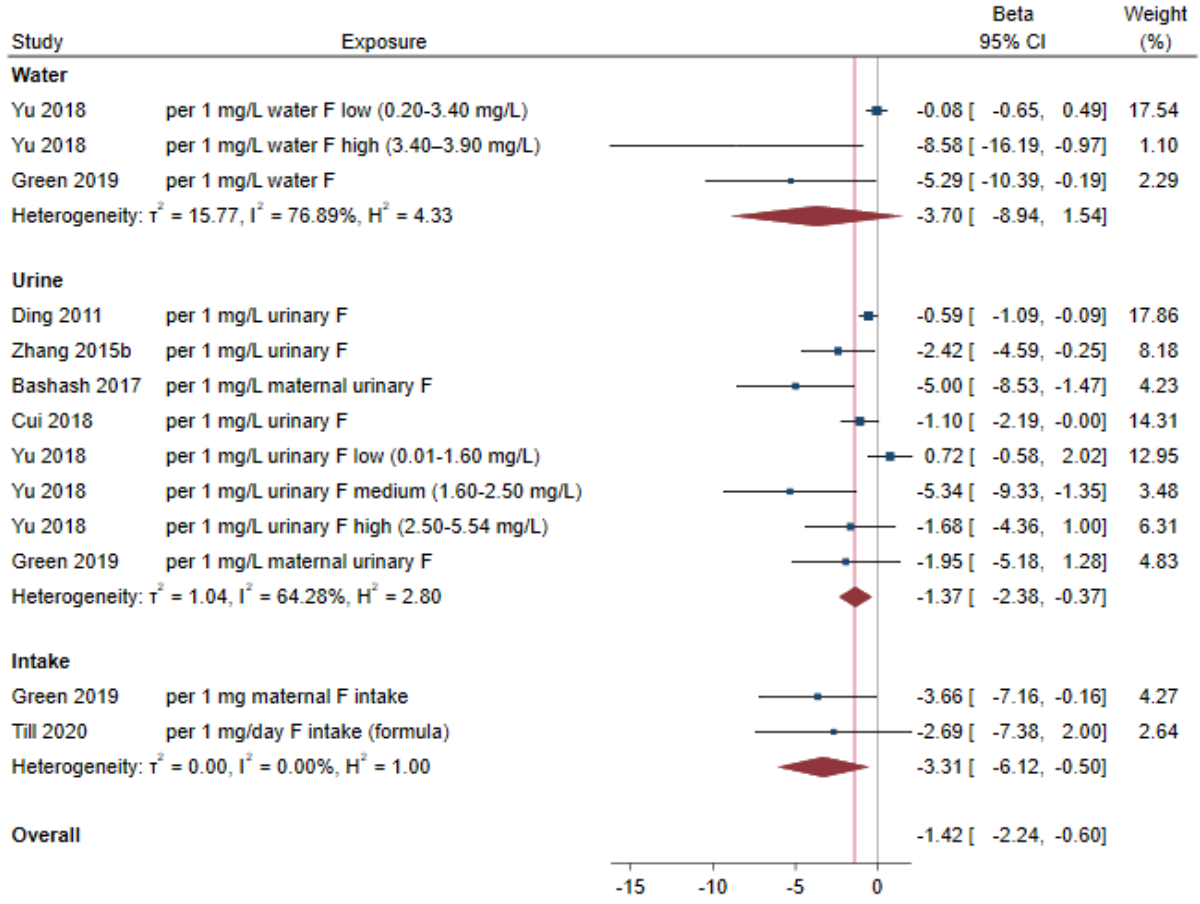


Figure B-5. Association Between Fluoride Exposure and IQ Scores in Children for Individual-level Exposures [No Pooling for Yu *et al.* (2018)]

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analyses. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Sensitivity Analysis for Individual-level Studies: Using Wang *et al.* (2020b) Versus Yu *et al.* (2018)

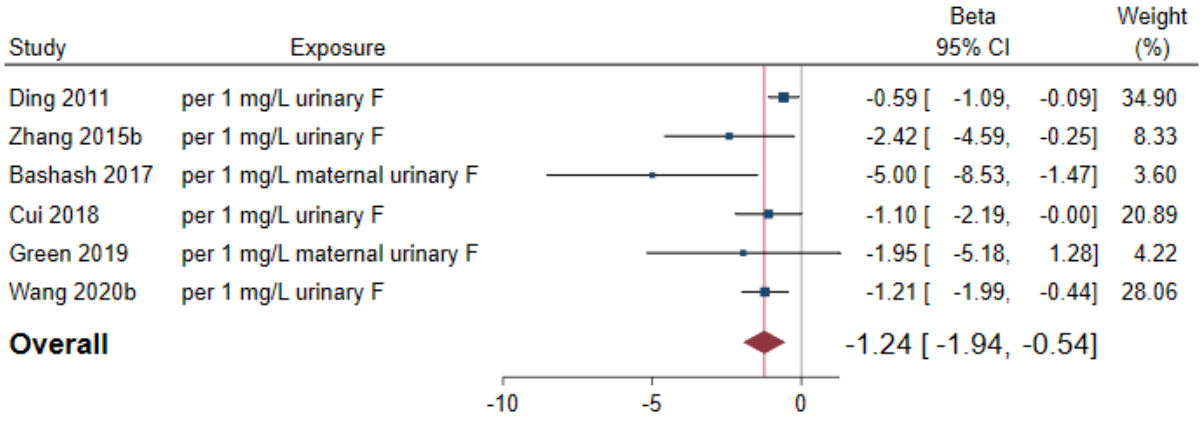


Figure B-6. Association Between Fluoride Exposure and IQ Scores in Children for Individual-level Exposures [Using Wang *et al.* (2020b) Estimates in Place of Yu *et al.* (2018) Estimates]

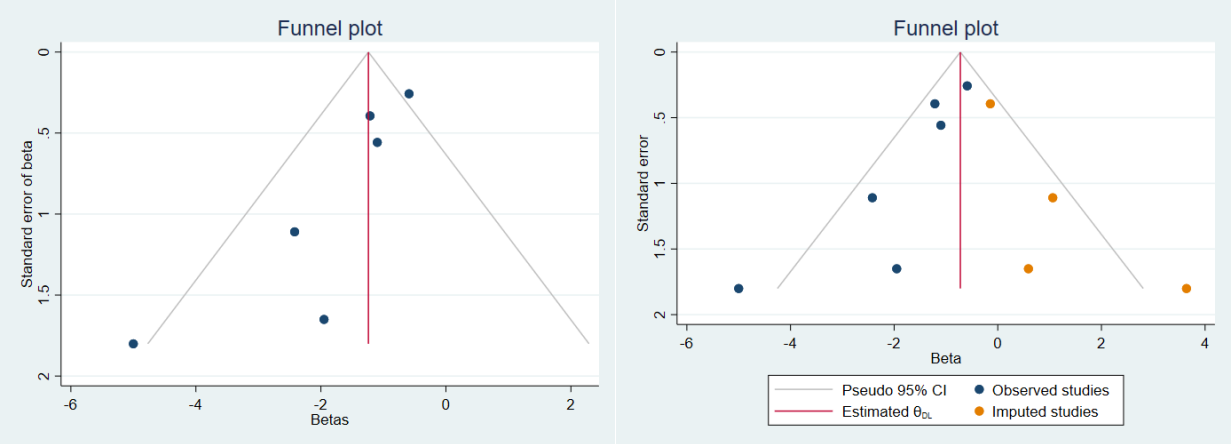


Figure B-7. Funnel Plot of Studies with Individual-level Exposures and Filled-in Funnel Plot to Eliminate Publication Bias [Using Wang *et al.* (2020b) Estimates in Place of Yu *et al.* (2018) Estimates]

Right panel shows funnel plot filled in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.97
      SE of beta1 =     0.712
                z =     -2.77
      Prob > |z| =    0.0056
```

Figure B-8. Test for Publication Bias for Studies with Individual-level Exposures [Using Wang *et al.* (2020b) Estimates in Place of Yu *et al.* (2018) Estimates]

```
Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration                          Number of studies =      9
  Model: Random-effects                     observed =      6
 Method: DerSimonian-Laird                  imputed =      3

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Beta	[95% Conf.	Interval]
Observed	-1.240	-1.939	-0.540
Observed + Imputed	-0.900	-1.676	-0.123

```
Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration                          Number of studies =     10
  Model: Random-effects                     observed =      6
 Method: DerSimonian-Laird                  imputed =      4

Pooling
  Model: Random-effects
 Method: DerSimonian-Laird
```

Studies	Beta	[95% Conf.	Interval]
Observed	-1.240	-1.939	-0.540
Observed + Imputed	-0.722	-1.411	-0.034

Figure B-9. Trim-and-fill Analysis for Studies with Individual-level Exposures [Using Wang *et al.* (2020b) Estimates in Place of Yu *et al.* (2018) Estimates]

Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

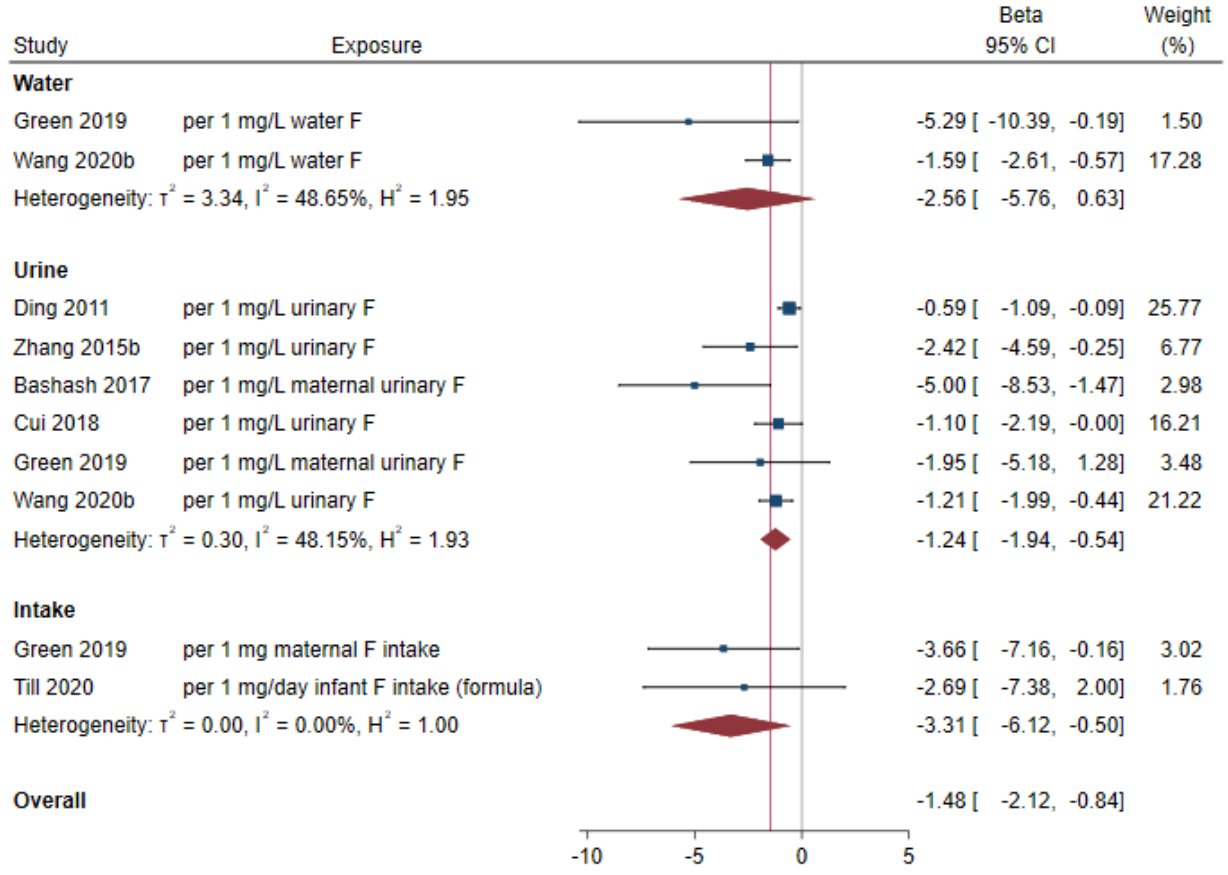


Figure B-10. Association Between Fluoride Exposure and IQ Scores in Children by Exposure for Individual-level Exposures [Using Wang *et al.* (2020b) Estimates in Place of Yu *et al.* (2018) Estimates]

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analysis. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Sensitivity Analysis for Individual-level Studies: Using Till *et al.* (2020) Versus Green *et al.* (2019)

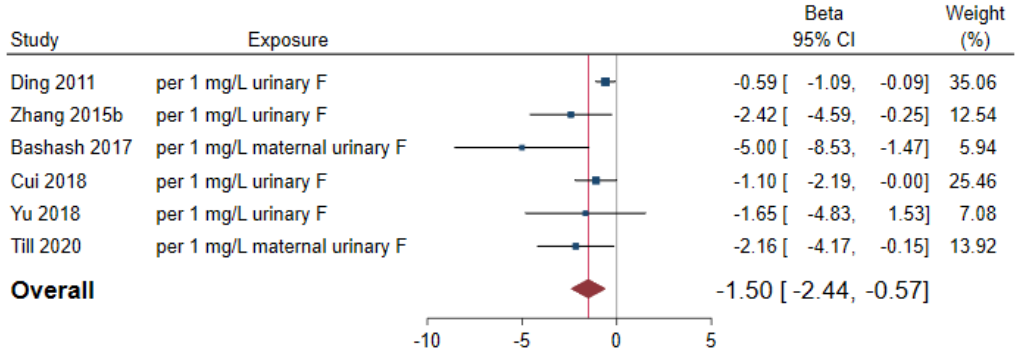


Figure B-11. Association Between Fluoride Exposure and IQ Scores in Children for Individual-level Exposures [Using Till *et al.* (2020) Estimates in Place of Green *et al.* (2019) Estimates]

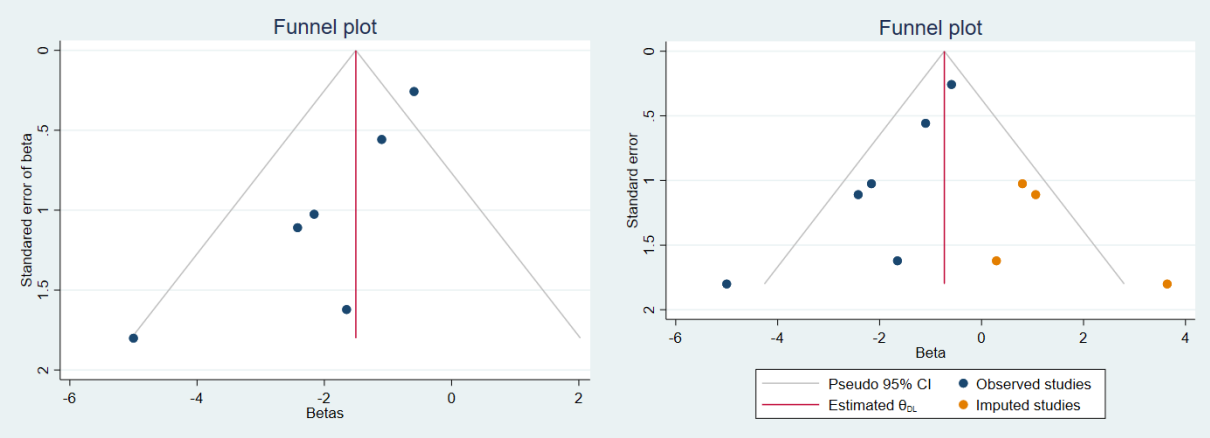


Figure B-12. Funnel Plot of Included Studies with Individual-level Exposures and Filled-in Funnel Plot to Eliminate Publication Bias [Using Till *et al.* (2020) Estimates in Place of Green *et al.* (2019) Estimates]

Right panel shows funnel plot filled in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.94
      SE of beta1 =     0.655
                z =     -2.96
      Prob > |z| =    0.0031
```

Figure B-13. Test for Publication Bias for Studies with Individual-level Exposures [Using Till *et al.* (2020) Estimates in Place of Green *et al.* (2019) Estimates]

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration — Number of studies = 9
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 3

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.504	-2.439	-0.570
Observed + Imputed	-0.808	-1.774	0.159

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration — Number of studies = 10
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 4

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.504	-2.439	-0.570
Observed + Imputed	-0.730	-1.649	0.188

Figure B-14. Trim-and-fill Analysis for Studies with Individual-level Exposures [Using Till *et al.* (2020) Estimates in Place of Green *et al.* (2019) Estimates]

Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

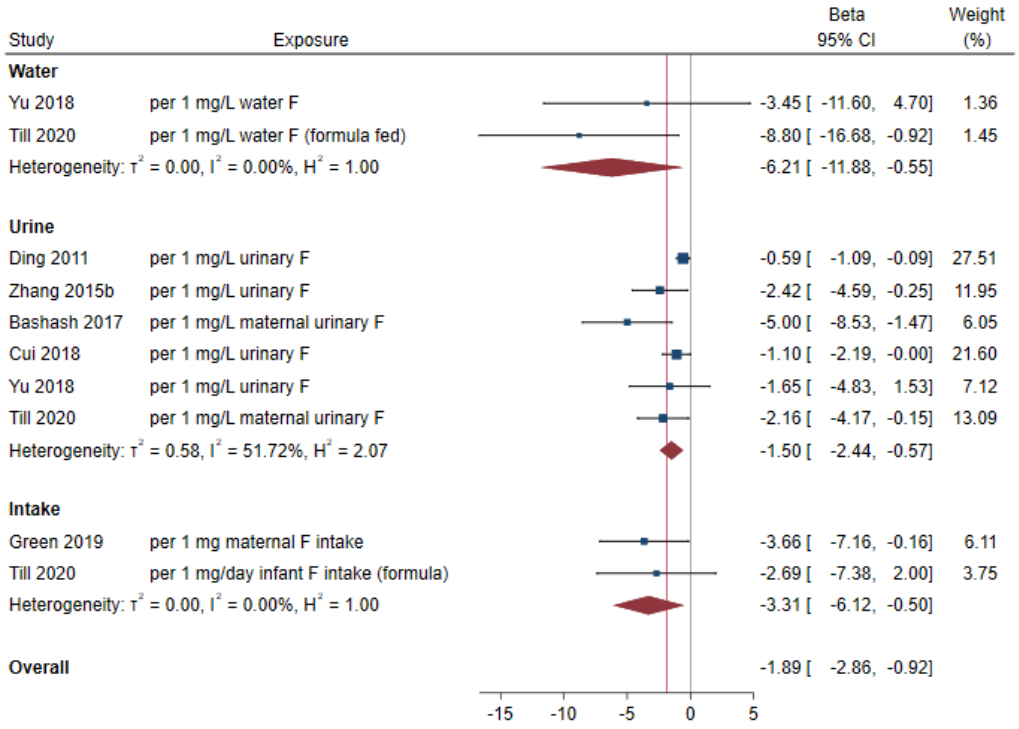


Figure B-15. Association Between Fluoride Exposure and IQ Scores in Children by Exposure Type for Individual-level Exposures [Using Till *et al.* (2020) Estimates in Place of Green *et al.* (2019) Estimates]

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analysis. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Sensitivity Analysis for Individual-level Studies: Verbal IQ

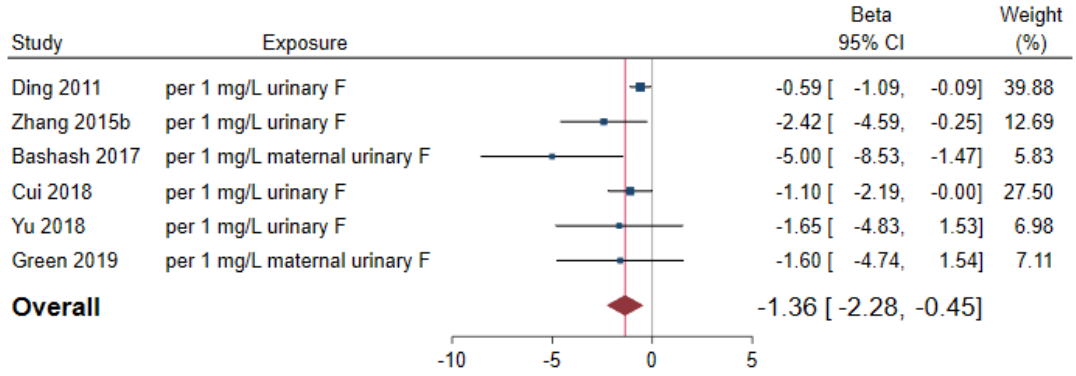


Figure B-16. Association between Fluoride Exposure and IQ Scores in Children Using Verbal IQ Score for Green *et al.* (2019)

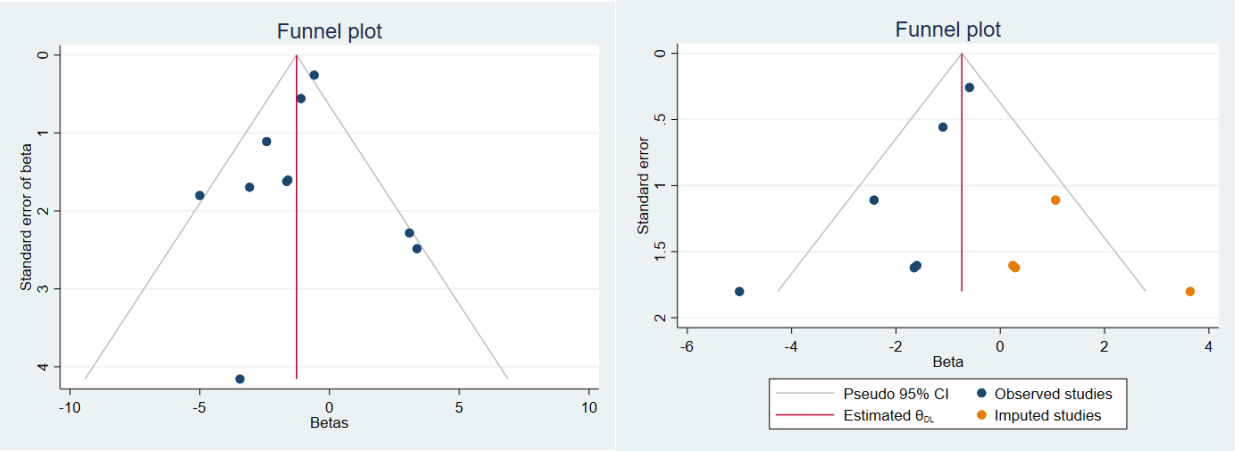


Figure B-17. Funnel Plot of Included Studies with Individual-level Exposures and Filled-in Funnel Plot to Eliminate Publication Bias [Using Verbal IQ Score for Green *et al.* (2019)]

Right panel shows funnel plot filled in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.62
      SE of beta1 =     0.629
                z =     -2.58
      Prob > |z| =    0.0098
```

Figure B-18. Test for Publication Bias for Studies with Individual-level Exposures [Using Verbal IQ Score for Green *et al.* (2019)]

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration — Number of studies = 9
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 3

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.365	-2.278	-0.452
Observed + Imputed	-0.814	-1.788	0.159

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration — Number of studies = 10
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 4

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.365	-2.278	-0.452
Observed + Imputed	-0.737	-1.654	0.180

Figure B-19. Trim-and-fill Analysis for Studies with Individual-level Exposures [Using Verbal IQ Score for Green *et al.* (2019)]

Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

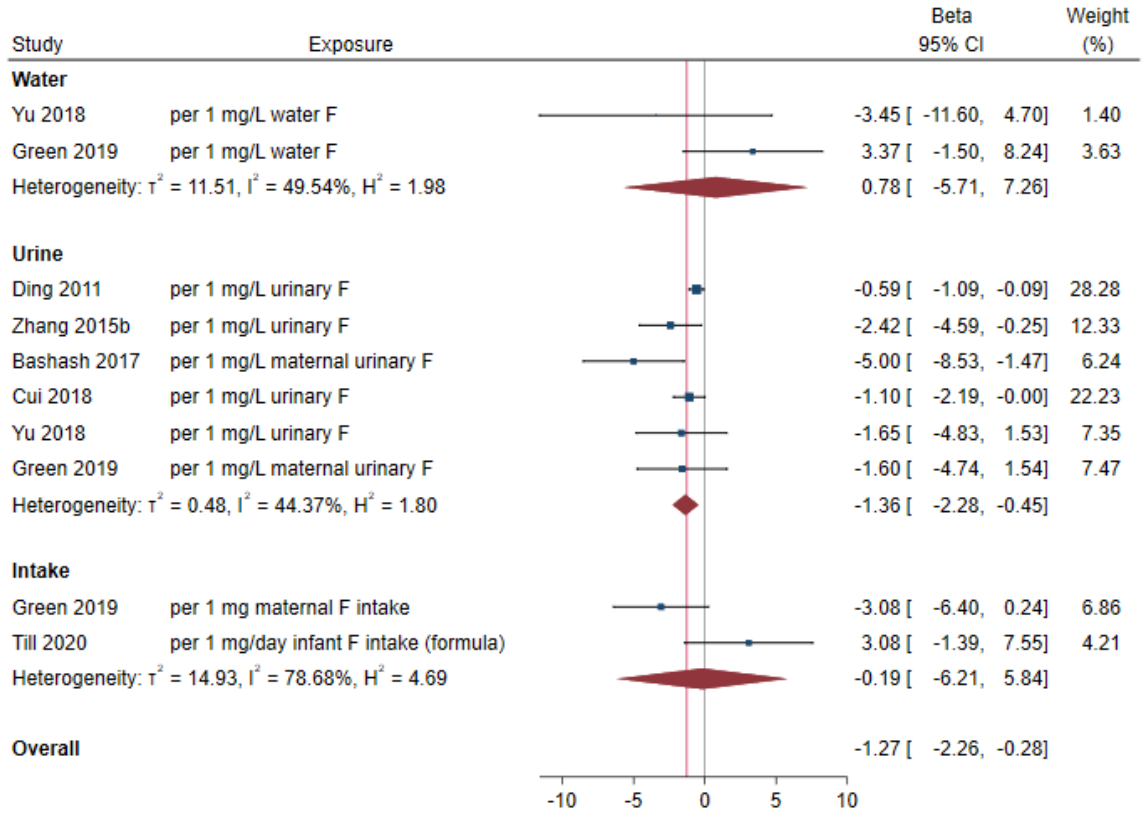


Figure B-20. Association Between Fluoride Exposure and IQ Scores in Children by Exposure Type Using Verbal IQ Score for Green *et al.* (2019)

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analysis. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Sensitivity Analysis for Individual-level Studies: Performance IQ

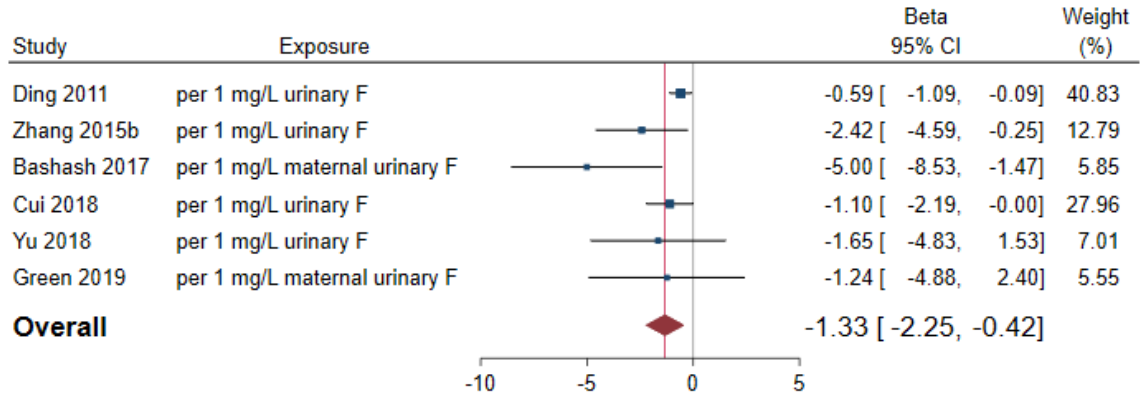


Figure B-21. Association Between Fluoride Exposure and IQ Scores in Children Using Performance IQ Score for Green *et al.* (2019)

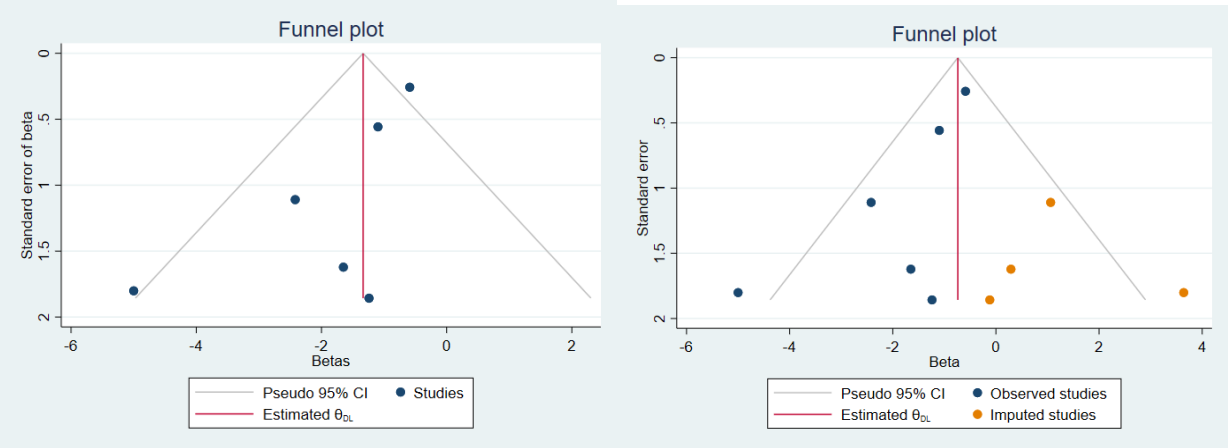


Figure B-22. Funnel Plot of Included Studies with Individual-level Exposures and Filled-in Funnel Plot to Eliminate Publication Bias [Using Performance IQ Score for Green *et al.* (2019)]

Right panel shows funnel plot filled in to the right using a run estimator. Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.52
      SE of beta1 =     0.623
                z =     -2.44
      Prob > |z| =    0.0147
```

Figure B-23. Test for Publication Bias for Studies with Individual-level Exposures [Using Performance IQ Score for Green *et al.* (2019)]

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration — Number of studies = 9
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 3

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.335	-2.248	-0.421
Observed + Imputed	-0.777	-1.759	0.204

Nonparametric trim-and-fill analysis of publication bias
Run estimator, imputing on the right

Iteration — Number of studies = 10
Model: Random-effects — observed = 6
Method: DerSimonian-Laird — imputed = 4

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.335	-2.248	-0.421
Observed + Imputed	-0.739	-1.662	0.184

Figure B-24. Trim-and-fill Analysis for Studies with Individual-level Exposures [Using Performance IQ Score for Green *et al.* (2019)]

Filling in to the left using a linear or a run estimator showed no change in the pooled effect estimate.

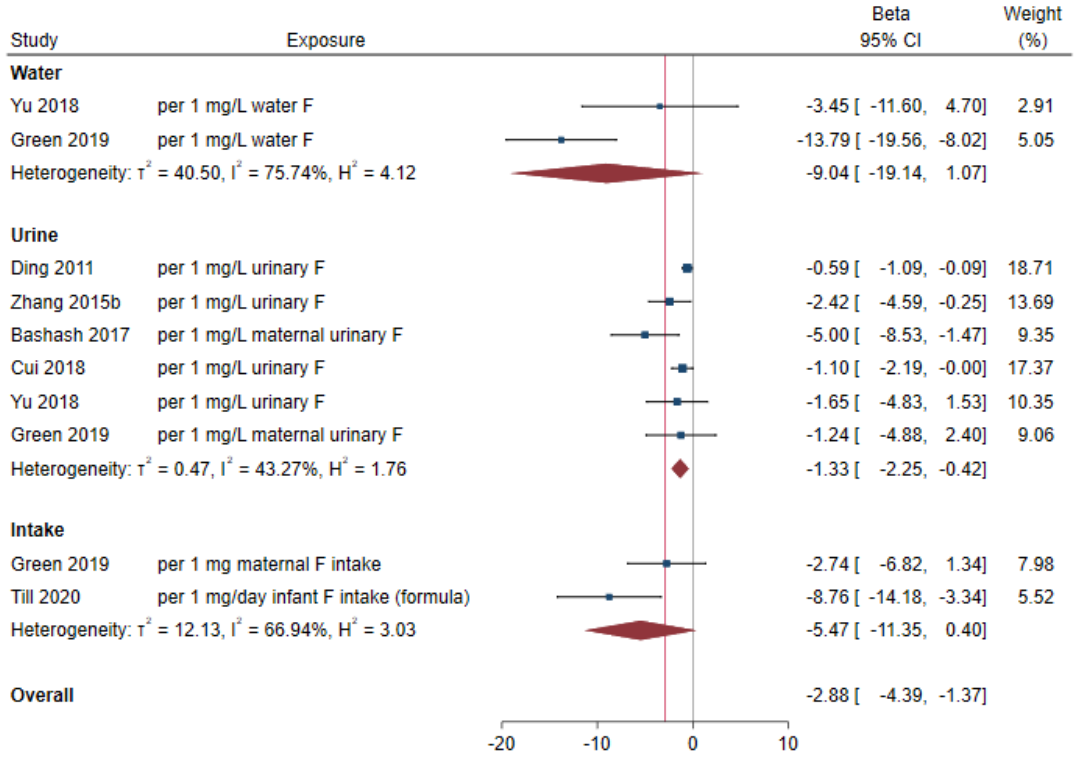


Figure B-25. Association Between Fluoride Exposure and IQ Scores in Children by Exposure Type Using Performance IQ Score for Green *et al.* (2019)

Note: The analysis for publication bias for studies with fluoride in urine is identical to the overall analysis. The analyses for publication bias for studies with fluoride in water or fluoride intake rely on two studies each and are not shown.

Sensitivity Analysis for Individual-level Studies: Excluding Cui *et al.* (2018)

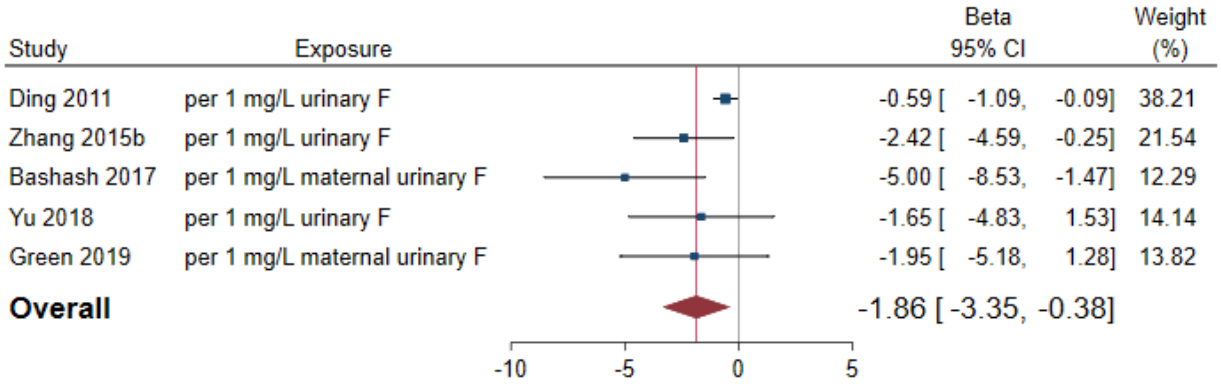


Figure B-26. Association Between Fluoride Exposure and IQ Scores in Children [Excluding Cui *et al.* (2018)]

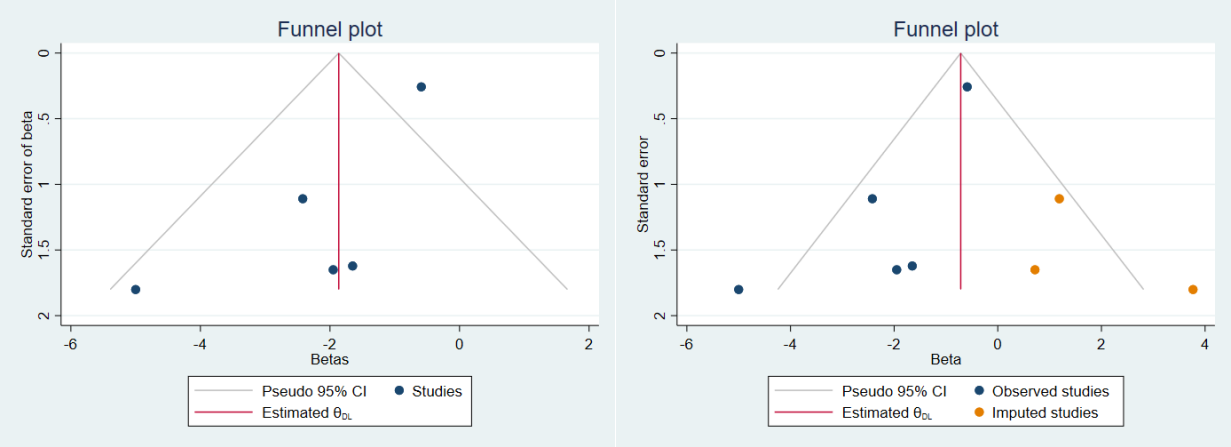


Figure B-27. Funnel Plot and Filled-in Funnel Plot to Eliminate Publication Bias of Included Studies with Individual-level Exposures [Excluding Cui *et al.* (2018)]

Right panel shows filling in to the right using a linear estimator. Filling in to the right using a run estimator or to the left using a linear or run estimator showed no change in the pooled effect estimate.

```
Regression-based Egger test for small-study effects
Random-effects model
Method: DerSimonian-Laird

H0: beta1 = 0; no small-study effects
            beta1 =     -1.68
      SE of beta1 =     0.638
                z =     -2.63
      Prob > |z| =    0.0085
```

Figure B-28. Test for Publication Bias for Studies with Individual-level Exposures [Excluding Cui *et al.* (2018)]

Nonparametric trim-and-fill analysis of publication bias
Linear estimator, imputing on the right

Iteration — Number of studies = 8
Model: Random-effects — observed = 5
Method: DerSimonian-Laird — imputed = 3

Pooling
Model: Random-effects
Method: DerSimonian-Laird

Studies	Beta	[95% Conf.	Interval]
Observed	-1.864	-3.351	-0.377
Observed + Imputed	-0.715	-2.056	0.625

Figure B-29. Trim-and-fill Analysis for Studies with Individual-level Exposures [Excluding Cui *et al.* (2018)]

Filling in to the right using a run estimator or to the left using a linear or run estimator showed no change in the pooled effect estimate.