DRAFT
NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Systematic Review

NTP Monograph 08

September 2022

National Toxicology Program
Public Health Service
U.S. Department of Health and Human Services
ISSN: 2378-5144

Research Triangle Park, North Carolina, USA

Foreword

The National Toxicology Program (NTP), established in 1978, is an interagency collaboration 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 this virtual program is administratively located. NTP's work focuses on 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.

Literature-based evaluations are one means by which NTP assesses 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 be used to make informed decisions about whether exposure to a substance may be of concern for human health.

These health effects evaluations follow prespecified 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."[†] 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.

Systematic review procedures are not algorithms, and the methods require scientific judgments. 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 judgments. 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.

[†]OHAT is the abbreviation for Office of Health Assessment and Translation, which has become the Health Assessment and Translation group in the Integrative Health Assessment Branch of the Division of the National Toxicology Program at the National Institute of Environmental Health Sciences.

Table of Contents

Tables

Figures

About This Review

National Toxicology Program[1]

[1]Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Collaborators

Kyla W. Taylor, John R. Bucher, Robyn B. Blain, Christopher A. Sibrizzi, Pamela A. Hartman, Kristen Magnuson, Sorina E. Eftim, and Andrew A. Rooney

Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA
Developed protocol and conducted systematic review through oversight and review of literature search, screening, data extraction, risk-of-bias assessment, and meta-analysis; wrote and edited monograph
Kyla W. Taylor, Ph.D., Project Lead
John R. Bucher, Ph.D.
Andrew A. Rooney, Ph.D.

ICF, Fairfax, Virginia, USA
Conducted literature screening, data extraction, risk-of-bias assessment, and meta-analysis; reviewed data, results, and analyses; wrote, edited, and formatted monograph; provided database and HAWC support
Robyn B. Blain, Ph.D.
Sorina E. Eftim, Ph.D.
Pamela A. Hartman, M.E.M.
Kristen Magnuson, M.E.S.M., Deputy Work Assignment Manager
Christopher A. Sibrizzi, M.P.H., Lead Work Assignment Manager

Commented [A1]: This section provides a new list of the names of all team members along with a description of tasks to which they contributed, follows our final document formatting approach, and is responsive to the [redacted] Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph page 11): Page 13: Should consider adding team member initials to their roles in the review.

Contributors

Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA
Provided oversight for external peer review
Mary S. Wolfe, Ph.D.

Provided literature search, screening, Distiller and HAWC support
Vickie R. Walker, B.S.

Reviewed meta-analysis protocol
Suril Mehta, Dr.P.H.

Conducted a technical review of the draft monograph
Freya Kamel, Ph.D. (Retired)

Conducted literature screening
Kris Thayer, Ph.D. (currently Chemical and Pollution Assessment Division, United States Environmental Protection Agency, Washington, D.C., USA)

Division of Extramural Research and Training, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA
Conducted literature screening
Heather King, Ph.D.

Kelly Government Services, Research Triangle Park, North Carolina, USA
Supported external peer review
Elizabeth A. Maull, Ph.D. (retired from NIEHS, Research Triangle Park, North Carolina, USA)

ICF, Fairfax, Virginia, USA
Provided contract oversight
David F. Burch, M.E.M., Principal Investigator
Cara Henning, Ph.D.
Jessica A. Wignall, M.S.P.H.

Reviewed data, results, and analyses
William Mendez, Ph.D.

Reviewed data, results, and analyses, conducted literature screening and risk-of-bias assessment, and provided database and HAWC support
Anna Engstrom, Ph.D.

Provided database and HAWC support
Steven Black, M.P.H.
Natalie Blanton, M.P.H.
Kaitlin Geary, B.S.
Courtney Lemeris, B.A.
Rachel McGill, B.S.
Daniel Peters, B.A.
Alessandria Schumacher, B.A.
Anna Stamatogiannakis, B.S.
River Williams, B.S.

Conducted literature screening
Carlye A. Austin, Ph.D.
Ryan Cronk, Ph.D.
Lauren Fitzharris, M.P.H.
Samantha Goodman, M.S.
Angela Hensel, B.A.
Melinda Hoang, M.P.H.
Allison Killius, B.S.
Alex Kliminsky, M.S.
Shannon McGinnis, Ph.D.

Devon Morgan, M.S.
Joei Robertson, M.Sc.
Courtney Rosenthal, M.S.
Amanda Ross, M.Ed.
Karen E. Setty, Ph.D.
Kelly Shipkowski, Ph.D.
Raquel A. Silva, Ph.D.
Parnian Soleymani, M.S.

Conducted literature screening and data extraction
Katie Duke, Ph.D.
Susan Goldhaber, M.P.H.
Revathi Muralidharan, B.S.
Jennifer Seed, Ph.D.
Codi Sharp, M.P.H.

Conducted literature screening, data extraction, and risk-of-bias assessment
Johanna Rochester, Ph.D.

Conducted literature screening and risk-of-bias assessment
Pam Ross, M.S.P.H.

Conducted data extraction
Robert Shin, M.H.S.

Conducted data extraction and risk-of-bias assessment
Tao Hong, Ph.D.
Alex Lindahl, M.P.H.
Samantha J. Snow, Ph.D., DABT

Conducted risk-of-bias assessment
Louise Assem, Ph.D.
Kristin Bornstein, Ph.D.
Jennifer Hsieh, M.S.P.H.
Yi Lu, Ph.D.

Conducted risk-of-bias assessment and reviewed data to support meta-analysis
Alexandra Goldstone, M.P.H.

Reviewed data to support meta-analysis
Jonathan Cohen, Ph.D.
Hannah Eglinton, B.A.
Cynthia J. Lin, Ph.D.

Supported meta-analysis visualizations
Rebecca Gray, M.P.H.

Coordinated peer review
Canden N. Byrd, B.S.

Designed and executed literature searches, and managed references
Michelle Cawley, M.A., M.L.S.
Jeremy S. Frye, M.S.L.S.
Nicole Vetter, M.S.L.S.

Edited and formatted monograph
Penelope Kellar, M.S.
Whitney Mitchell, B.S.
Kevin O'Donovan, B.A.

Integrated Laboratory Systems, LLC, Research Triangle Park, North Carolina, USA
Conducted a technical review of the draft monograph
Cynthia J. Willson, D.V.M., Ph.D., DACVP

Boston University, Boston, Massachusetts, USA
Reviewed the systematic review protocol
Thomas Webster, Ph.D.

Brown University, Providence, Rhode Island, USA
Reviewed the systematic review protocol
Joseph Braun, Ph.D.

Columbia University, New York, New York, USA
Reviewed the systematic review protocol
Gail Wasserman, Ph.D.

Health Research Board, Dublin, Ireland
Reviewed the systematic review protocol
Marie Sutton, Ph.D.

University of Massachusetts, Amherst, Massachusetts, USA
Reviewed the systematic review protocol
Thomas Zoeller, Ph.D.

University of Colorado, Denver, Colorado, USA
Reviewed meta-analysis protocol
Tianjing Li, Ph.D.

Peer Review

The National Toxicology Program (NTP) conducted a peer review of the draft *NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Systematic Review* by letter in December 2021. Reviewer selection and document review followed established NTP practices. The reviewers were charged to:

(1) Comment on the technical accuracy and whether the draft *NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects* is clearly stated and objectively presented.

(2) Determine whether the scientific evidence supports the NTP's confidence ratings for the bodies of evidence regarding neurodevelopmental and cognitive health effects associated with exposure to fluoride.

NTP carefully considered reviewer comments in finalizing this monograph.

Peer Reviewers

Ethan Balk, M.D., M.P.H.
Associate Professor (Research)
Brown University School of Public Health
Providence, Rhode Island, USA

Pam Factor-Litvak, Ph.D.
Professor of Epidemiology
Columbia University Medical Center
New York, New York, USA

Erin Haynes, Dr.P.H.
Kurt W. Deuschle Professor in Preventive Medicine and Environmental Health
University of Kentucky College of Public Health
Lexington, Kentucky, USA

Julie Obbagy, Ph.D., RD
Nutritionist, Center for Nutrition Policy and Promotion
Office of Nutrition Guidance and Analysis
U.S. Department of Agriculture
Medfield, Massachusetts, USA

Heather Volk, Ph.D.
Associate Professor
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland, USA

Publication Details

Publisher: National Toxicology Program

Publishing Location: Research Triangle Park, NC

ISSN: 2378-5144

DOI: https://ntp.niehs.nih.gov/go/785076

Report Series: NTP Monograph Series

Report Series Number: 08

Official citation: National Toxicology Program (NTP). 2022. NTP monograph on the state of the science concerning fluoride exposure and neurodevelopmental and cognitive health effects: a systematic review. Research Triangle Park, NC: National Toxicology Program. NTP Monograph 08.

Acknowledgements

This work was supported by the Intramural Research Program (ES103316, ES103317) at the National Institute of Environmental Health Sciences, National Institutes of Health and performed for the National Toxicology Program, Public Health Service, U.S. Department of Health and Human Services under contract GS00Q14OADU417 (Order No. HHSN273201600015U).

Conflict of Interest

Individuals who reviewed the systematic review protocol or meta-analysis protocol, conducted a technical review of the draft monograph, or served on the peer review panel have certified that they have no known real or apparent conflict of interest related to fluoride exposure or neurodevelopmental and cognitive health effects.

Abstract

Background: Fluoride is a common exposure in our environment that comes from a variety of sources and is widely promoted for its dental and overall oral health benefits. 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. The evidence reviewed at that time was from dental and skeletal fluorosis-endemic regions of China. Since the NRC evaluation, the number and location of studies examining cognitive and neurobehavioral effects of fluoride in humans have grown considerably, including several recent North American prospective cohort studies evaluating prenatal fluoride exposure.

In 2016, the National Toxicology Program (NTP) published a systematic review of the evidence from experimental animal studies on the effects of fluoride on learning and memory. That systematic review found a low-to-moderate level of evidence that deficits in learning and memory occur in non-human mammals exposed to fluoride.

Objective: To conduct a systematic review of the human, experimental animal, and mechanistic literature to evaluate the extent and quality of the evidence linking fluoride exposure to neurodevelopmental and cognitive effects in humans.

Method: A systematic review protocol was developed and utilized following the standardized OHAT systematic review approach for conducting literature-based health assessments. This monograph presents the current state of evidence associating fluoride exposure with cognitive or neurodevelopmental health effects and incorporated predefined assessments of study quality and confidence levels. Benefits of fluoride with respect to oral health are not addressed in this monograph.

Results: The current bodies of experimental animal studies and human mechanistic evidence do not provide clarity on the association between fluoride exposure and cognitive or neurodevelopmental human health effects.

This systematic review identified studies that assessed the association between fluoride exposure and cognitive or neurodevelopmental effects in both adults and children, which were evaluated separately. In adults, only two high-quality cross-sectional studies examining cognitive effects were available. The literature in children was more extensive and was separated into studies assessing intelligence quotient (IQ) and studies assessing other cognitive or neurodevelopmental outcomes. Eight of nine high-quality studies examining other cognitive or neurodevelopmental outcomes reported associations with fluoride exposure. Seventy-two studies assessed the association between fluoride exposure and IQ in children. Nineteen of those studies were considered to be high quality; of these, 18 reported an association between higher fluoride exposure and lower IQ in children. The 18 studies, which include 3 prospective cohort studies and 15 cross-sectional studies, were conducted in 5 different countries. Forty-six of the 53 low-quality studies in children also found evidence of an association between higher fluoride exposure and lower IQ in children.

Discussion: Existing animal studies provide little insight into the question of whether fluoride exposure affects IQ. In addition, studies that evaluated fluoride exposure and mechanistic data in humans were too heterogenous and limited in number to make any determination on biological plausibility. The body of evidence from studies in adults is also limited and provides low

confidence that fluoride exposure is associated with adverse effects on adult cognition. There is, however, a large body of evidence on IQ effects in children. There is also some evidence that fluoride exposure is associated with other neurodevelopmental and cognitive effects in children; although, because of the heterogeneity of the outcomes, there is low confidence in the literature for these other effects. This review finds, with moderate confidence, that higher fluoride exposure (e.g., represented by populations whose total fluoride exposure approximates or exceeds the World Health Organization Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride) is consistently associated with lower IQ in children. More studies are needed to fully understand the potential for lower fluoride exposure to affect children's IQ.

Commented [A2]: The following two sentences reflect revisions to clarify that our moderate confidence conclusion is primarily based on studies with total fluoride exposure that approximates or exceeds what is generally associated with consumption of optimally fluoridated water in the United States, in response to the [redacted] comments below; see DocA1_Monograph and DocB1_Monograph, respectively, for detailed response:

Reviewer comment (DocA1_Monograph, page 2): Recommendation: [redacted] requests NTP include a statement in the systematic review abstract and full text, as well as the meta-analysis, like that found in the 2020 draft monograph: "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) that can be evaluated for dose response, effects on cognitive neurodevelopment are inconsistent, and therefore unclear."

Reviewer comment (DocB1_Monograph, page 2): It would be helpful if the Abstract was clear in the Discussion that the conclusion about effects on IQ in children was derived from high human exposures (higher than US exposures) without getting into more hazard conclusions or assessments.

Commented [A3]: This sentence and similar sentences throughout the monograph reflect revisions to clarify that our moderate confidence conclusion is primarily based on studies with total fluoride exposure that approximates or exceeds what is generally associated with consumption of optimally fluoridated water in the United States, in response to the [redacted] comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 6): Dose response: The authors correctly point out that many of the studies dealt with exposure levels considered relatively high, at least relative to the EPA drinking water standard and secondary standard. Further, in the group of studies considered to be low risk of bias, exposure was generally considered either on an arithmetic – but sometimes on a logarithmic scale (if quantitative), or based on a dichotomous variable of fluorosis, a manifestation of continuous high exposure, or whether study participants lived in an area known to have high levels of exposure. Thus, the conclusion of moderate confidence that fluoride is associated with deficits in IQ scores in children needs to be couched for these higher exposure levels, as there are few studies that provide evidence of this for exposures in the low range. This is not to say that there is no association at these lower levels, there may very well be an association; just that these results cannot be generalized to lower levels of exposure. This is true with other neurotoxins as well, for example, we know that the associations between lead and IQ scores is even steeper at the lower levels of exposure, but early studies where exposure was high were not able to discern those associations.

Preface

The National Toxicology Program (NTP) conducted a systematic review of the published scientific literature because of public concern regarding the potential association between fluoride exposure and adverse neurodevelopmental and cognitive health effects.

NTP initially published a systematic review of the experimental animal literature in 2016 that was subsequently expanded to include human epidemiological studies, mechanistic studies, and newer experimental animal literature. Because of the high public interest in fluoride's benefits and potential risks, NTP asked the National Academies of Sciences, Engineering, and Medicine (NASEM) to conduct an independent evaluation of the draft *NTP Monograph on Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects* (2019 draft monograph dated September 6, 2019) and the revised draft (2020 draft monograph dated September 16, 2020), which addressed the NASEM committee's recommendations for improvement. The NASEM committee determined that, "Overall the revised monograph seems to include a wealth of evidence and a number of evaluations that support its main conclusion, but the monograph falls short of providing a clear and convincing argument that supports its assessments…." Thus, NTP has removed the hazard assessment step and retitled this systematic review of fluoride exposure and neurodevelopmental and cognitive health effects as a "state-of-the-science" document to indicate the change. This state-of-the-science document does not include the meta-analysis of epidemiological studies or hazard conclusions found in previous draft monographs; however, it provides a comprehensive and current assessment of the scientific literature on fluoride as an important resource to inform safe and appropriate use.

NTP has responded to the NASEM committee's comments on the revised draft (September 16, 2020) in a separate document (*Sup01_Monograph_NASEM_Feb_2021.pdf*) and revised relevant sections of this monograph.

Introduction

Fluoride is a common exposure in our environment from a variety of sources and is widely promoted for its dental and overall oral health benefits. Approximately 67% of the U.S. population receives fluoridated water through a community 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 that 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. For many years, most fluoridated community water systems used fluoride concentrations ranging from 0.8 to 1.2 milligrams/liter (mg/L) (US DHHS 2015). For community water systems that add fluoride, PHS now recommends a fluoride concentration of 0.7 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 (MCL), is 4.0 mg/L. This level is the maximum amount of fluoride contamination (naturally occurring, not from water fluoridation) that is allowed in water from public water systems and 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 of fluoride, 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 teeth. Although the secondary standard is not enforceable, EPA requires that public water systems notify the public if and when average fluoride levels exceed 2.0 mg/L (NRC 2006). The World Health Organization (WHO) set a safe water guideline of 1.5 mg/L of fluoride in drinking water (first established in 1984 and reaffirmed in 1993 and 2011), which is recommended to protect against increasing risk of dental and skeletal fluorosis (WHO 2017).

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) (CDC Division of Oral Health 2020).

~~Commonly cited h~~Health concerns ~~related~~ cited in relation to fluoride are ~~bone fractures and~~ skeletal fluorosis, lower intelligence quotient (IQ) and other neurological effects, cancer, and

Commented [A4]: The organization of the Introduction section reflects revisions to have the uses of and exposure sources to fluoride be the first topics covered, in response to the [redacted] comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 4): The scientific information presented appears technically correct and objectively presented. A few suggestions are noted below to improve clarity.
The background section could be reorganized for clarity and flow. It might be beneficial to begin the abstract and background with the pervasive use of fluoride in drinking water followed by a brief statement of the benefits. The benefits of fluoride in water has not been articulated. The benefits only need a sentence or two. The background appears to be more of a justification for the report rather than a true background of the evidence leading to the study/report.

Commented [A5]: Change made in response to the [redacted] editorial comment below; see DocB2_Monograph for detailed response:

Reviewer comment (DocB2_Monograph, page 4): ~~Commonly cited health~~ Health concerns ~~related~~ cited in relation to fluoride are bone fractures and skeletal fluorosis, lower intelligence quotient (IQ) and other neurological effects, cancer, and endocrine disruption." [Text in red text inserted and red-strikethrough font deleted by the [redacted].]

Commented [A6]: Reference to 'bone fractures' removed in response to [redacted] comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 6): This statement is inflammatory. It is not a reflection of the current state of the science on this issue. However, these assertions that have been made by the Fluoride Action Network and are not evidenced-based. Ref: Osteoporos Int. 2008 Mar;19(3):257-68. Epub 2007 Aug 15.
Effects of treatment with fluoride on bone mineral density and fracture risk--a meta-analysis

P Vestergaard 1, N R Jorgensen, P Schwarz, L Mosekilde

Affiliations expand PMID: 17701094

endocrine disruption. Effects on neurological function, endocrine function (e.g., thyroid,[1] 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. The NRC report concluded that the Maximum Contaminant Level Goal (MCLG), 4 mg/L, should be lowered to protect against severe enamel fluorosis and reduce the risk of bone fractures associated with skeletal fluorosis (NRC 2006). Other than severe fluorosis, NRC did not find sufficient evidence of negative health effects at fluoride levels below 4 mg/L; however, it 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, including 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.

In 2016, the National Toxicology Program (NTP) 2016, NTP published a systematic review of the evidence from experimental animal studies on the potential effects of fluoride exposure on learning and memory (NTP 2016). That systematic review found a low-to-moderate level of evidence that deficits in learning and memory occur in experimental animals exposed to fluoride. Given these findings, NTP decided to conduct additional animal studies before carrying out this full systematic review and integrate human, animal, and potentially relevant mechanistic evidence in order to reach human health 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 monograph 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 health impacts based on time frame of exposure (i.e., during development or during adulthood). The evaluation of experimental animal studies in this monograph has been conducted separately from the 2016 experimental animal assessment; however, like the 2016 assessment, it 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.

A committee convened by the National Academies of Sciences, Engineering, and Medicine (NASEM) reviewed earlier drafts of this monograph (September 6, 2019, and September 16, 2020) (NASEM 2020; 2021). The current document incorporates changes stemming from those reviews, and responses to the 2020 review are available as

[1]The current review has evaluated the fluoride literature with an eye toward potential thyroid effects because a large literature base has accumulated examining the interaction of fluoride with iodine uptake by the thyroid gland and consequential effects on synthesis of thyroid hormones, which are recognized to play significant roles in neurodevelopment in utero and during early childhood. This literature, along with a detailed proposed mechanism of action, was recently reviewed by Waugh (2019).

Commented [A7]: This footnote was added to increase clarity regarding the purpose of examining published data on thyroid function, in response to similar comments from two [redacted] Reviewers (note one of them provided two separate comments on this issue); see DocG_Monograph and DocI_Monograph, respectively, for detailed responses:

Reviewer comment (DocG_Monograph, page 23): A clearer statement is needed up front, ideally in the Introduction about what topics were covered by full systematic review (which are a small subset of topics of interest) and why. It's very confusing to read through repeated descriptions of topics which are not being reviewed. As an example, it's disconcerting to repeatedly see that thyroid function is an outcome of interest (without an explanation as to why this is of interest to a review of neurodevelopmental and cognitive health effects) and then to come across the statement (page 13) that "Thyroid data were … not extracted." It's difficult to pick out and follow the reasoning for excluding most topics from full evaluation. The timing of and reasoning for the decisions to focus the systematic review on just "high quality" pediatric studies is unclear.

Reviewer comment (DocG_Monograph, page 23): Of note, the Introduction and Methods do not explain why thyroid function was evaluated. This was only (partially done) on page 63.

Reviewer comment (DocI_Monograph, page 6): Thyroid function isn't mentioned until the specific aims. It should be included in background along with other possible mechanisms, if known. It is unclear why thyroid function is being evaluated as the only mechanistic pathway. A figure or illustration depicting the theoretical pathway would be helpful.

Sup01_Monograph_NASEM_Feb_2021.pdf. See Appendix B, Table B-1 for a timeline of key activities contributing to this 2022 NTP monograph, including document review activities that have occurred since 2016.

Objective and Specific Aims

Objective

The overall objective of this evaluation was to undertake a systematic review to develop NTP human health hazard identification conclusions on the association between exposure to fluoride and neurodevelopmental and cognitive effects based on assessing levels of evidence from human and non-human animal studies with consideration of the degree of support from mechanistic data. However, the NASEM Committee's reviews (NASEM 2020; 2021) of the 2019 and 2020 drafts of the monograph indicated that, "Overall the revised monograph seems to include a wealth of evidence and a number of evaluations that support its main conclusion, but the monograph falls short of providing a clear and convincing argument that supports its assessments…." For this reason, our methods were revised to remove the hazard assessment step (i.e., the section "Integrate Evidence to Develop Hazard Identification Conclusions" and the associated section "Translate Confidence Ratings into Level of Evidence for Health Effect"). In addition, a meta-analysis of the epidemiological studies examining children's IQ in relation to fluoride exposure added to the 2020 draft in response to NASEM comments (NASEM 2020) ~~will be published separately~~ was removed for further refinement in preparation for a separate publication and is not part of this document.

Commented [A8]: Change made in response to the [redacted] comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, page 12): I would not presume that this will happen. Revise.

Therefore, the objective of this monograph is to undertake a systematic review of the literature concerning the association between fluoride exposure and neurodevelopmental and cognitive effects and to determine the level of confidence in that evidence. The assessment was based on evidence from human and non-human animal studies with consideration of mechanistic information.

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 neurobehavioral[2] function.
- Summarize the extent and types of health effects evidence available.

Commented [A9]: This footnote was added, and text here and throughout the monograph reflect revisions to change the term "neurologic" to "neurobehavioral" (or other appropriate text), in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 4): The term neurologic to refer to outcomes such as anxiety and aggression (and other neurobehavioral outcomes) is not quite correct. Neurologic would refer to outcomes such as tremor or other objective neurological signs. The more correct term would be neurobehavioral. I have marked some of this in the text.

[2]The specific aim in the protocol refers to "impaired neurological function"; however, it was changed to "impaired neurobehavior function" in this document to use more precise terminology. The overall aim from the protocol remained the same for this evaluation.

- Describe limitations of the systematic review, strengths and 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.

Depending on the extent and nature of the available evidence:

- Synthesize the evidence using a narrative approach.
- 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.

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 a 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 (OEHHA 2011; NRC 2006; 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.

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).[3] The protocol served as the complete set of methods followed for the conduct of this systematic review. The OHAT Handbook for Conducting a Literature-Based Health Assessment (http://ntp.niehs.nih.gov/go/38673) is a source of general systematic review methods that were selected and tailored in developing this protocol. Options in the OHAT handbook that were not specifically referred to in the protocol were not part of the methods for the systematic review.

A brief summary of the methods is presented below. Although the methods were revised to remove the hazard assessment step and meta-analysis from this document, the protocol was not further revised.

PECO Statements

PECO (**P**opulation, **E**xposure, **C**omparators and **O**utcomes) statements were developed as an aid to identify search terms and appropriate inclusion/exclusion criteria for addressing the overall research question (effects on neurodevelopmental or cognitive function and thyroid associated

[3]NTP conducts systematic reviews following prespecified 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 that supersede the methods in the OHAT Handbook.

Commented [A10]: Footnote 3 reflects revisions to clarify that the protocol describes all the methods used for this systematic review described in the monograph in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 3): Recommendation: NTP should specify the areas where they departed from the OHAT protocol.

Commented [A11]: The following two sentences reflect revisions to further clarify the role of the OHAT handbook and the stand-alone nature of the protocol, in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 2): See NASEM comment about the protocol on p. 4 of Response to Fluoride NASEM Letter 10.5.2021:"Although the statement clarifies the general role of the handbook, the committee finds that it does not address the committee's previous recommendation to set the expectation for how closely the process described in the handbook will be followed in the protocol and in the eventual systematic review. For example, the handbook section "Key Questions and Analytical Framework" that guides development of the population, exposure, comparator, and outcomes (PECO) statement is not included in the fluoride protocol or the revised monograph. As the committee recommended in its previous review, NTP should treat each systematic review protocol as a stand-alone document that contains all the information necessary for understanding of the planning and conduct of the review, and these expectations should be explicitly stated in the protocol. The committee did not find that revisions of the protocol adequately addressed this recommendation.

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 human studies, controlled exposure animal studies, and mechanistic/in vitro studies for evidence of neurodevelopmental or cognitive function and thyroid effects associated with fluoride exposure. 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 and 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) to evaluate the available information. To prioritize and consider available mechanistic data, the ~~Categories~~ 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 effects on learning and memory but to evaluate whether a plausible series of mechanistic events exists to support effects observed in the low-dose region (below approximate drinking-water-equivalent concentrations of 20 ppm for animal studies) that may strengthen a hazard conclusion if one is derived.

Commented [A12]: This sentence was revised clarify how mechanistic data were prioritized for consideration in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 5): Would it be possible to define what is meant by "Categories focused on were those with more robust data at levels of fluoride more relevant to human exposure"? Should this information be documented as part of the PECO? Was this an inclusion criteria, or just used in prioritizing or weighting the evidence in drawing conclusions?

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; Chemical Abstracts Service Registry Number [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 (e.g., exposure below detection levels) or exposed to lower levels of fluoride ~~(e.g., exposure below detection levels)~~[4]
Outcomes	Neurodevelopmental outcomes, including learning, memory, intelligence, other forms of cognitive behavior, other neurological/neurobehavioral[5] outcomes (e.g., anxiety, aggression, motor activity), and biochemical changes in the brain or nervous system tissue; measures of thyroid function, biochemical changes, or thyroid tissue pathology

Commented [A13]: Change made in response to the [redacted] Reviewer comment below and footnote #4 added to clarify populations exposed to lower levels of fluoride as the comparator; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 7): how would you determine "detection" levels if you are not measuring the dose of exposure? Not clear how comparable popltns *(sic)* of not exposed are equivalent or appropriate to use in lieu of "exposed to lower levels of fluoride"

[4]Note: The "(e.g., exposure below detection limits)" was moved after "populations not exposed to fluoride" to reflect how it was used in the literature search and elsewhere in this systematic review.

[5]The human PECO statement in the protocol refers to "neurological outcomes"; however, it was changed to "neurological/neurobehavioral outcomes" in this document to use more precise terminology for the outcomes included.

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/neurobehavioral[6] outcomes (e.g., anxiety, aggression, motor activity), and biochemical changes in the brain or nervous system tissue; measures of thyroid function, biochemical changes, or thyroid tissue pathology

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

[6]The animal PECO statement in the protocol refers to "neurological outcomes"; however, it was changed to "neurological/neurobehavioral outcomes" in this document to use more precise terminology for the outcomes included.

fluoride by reviewing Medical Subject Headings for relevant and appropriate neurobehavioral and thyroid-related terms and by extracting key neurobehavioral and thyroid-related health effects and developmental neurobehavioral terminology from reviews and a sample of relevant studies.[7] Combinations of relevant subject headings and keywords were subsequently identified. A test set of relevant studies was used to ensure the search terms retrieved 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 B; 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.

An additional search was conducted on May 1, 2020, where human epidemiological studies with primary neurodevelopmental or cognitive outcomes (learning, memory, and intelligence) were prioritized during screening. The review of the 2020 search results focused only on the human studies because they formed the basis of the confidence ratings (see Figure 1 for framework to assess confidence) and conclusions in the September 6, 2019, draft. A supplemental literature search of Chinese-language databases (described below) was also conducted. See Appendix B, Table B-1 for a timeline of key activities contributing to this 2022 NTP monograph, including information relevant to the timing of multiple literature searches.

Publications identified in these searches are categorized as "references identified through database searches" in Figure 2. Studies identified from other sources or manual review that satisfy the PECO criteria for inclusion ~~might impact conclusions~~ are considered under "references identified through other sources" in Figure 2. 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 experimental animal studies prior to 2015, the current assessment did not evaluate these studies 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

In order to identify non-English-language studies that might not appear in databases for the main literature search, additional searches were developed for non-English-language databases. No definitive guidance was found on the most comprehensive, highest quality, or otherwise most appropriate non-English-language databases for health studies of fluoride. Therefore, databases were chosen that identified non-English-language studies that were not captured in searches of databases from the main literature search—those previously identified from other resources (see the Searching Other Resources section below). Multiple non-English-language databases were

[7]The terms "study" and "publication" are used interchangeably in this document to refer to a published work drawn from an original body of research conducted on a defined population.

Commented [A14]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 8): what is the relationship of the 11 studies identified through these means proportional to those included from the database searches? it is concerning that there were 11 studies that were NOT identified in the database search that may have been important... does this represent literature/study selection bias?

Commented [A15]: This paragraph and the next reflect revisions to the monograph that describe the additional steps the review team took to translate and extract data from all non-English language studies identified from the Chinese database searches (including some that were not previously extracted), in response to the [redacted] Reviewer comments below; see DocG_Monograph, DocH_Monograph, and DocJ_Monograph, respectively, for detailed responses:

Reviewer comment (DocG_Monograph, page 4): Regarding the overall search (for all topics) in the Chinese databases (page 10), the strategy as described is unacceptable and flies in the face of the goal of systematic review. The statements "A primary goal of the screening of the newly-retrieved human references in the supplemental search of Chinese databases was to identify null, or no-effect, studies" and "Null studies that were identified were translated and included." A plain reading of these statements suggests a high degree of bias by the researchers such that evidence of an association (not null studies) were omitted. All studies should be found and included, regarding of findings.
That said, it may be that your aim was to identify studies missed due to reporting or publication bias. If that is the case, this should be stated as the primary goal. However, it might just be better to drop this concept altogether, since the purpose of searching the non-English databases was to capture studies that have not made it into (primarily Western) databases. This is sufficient explanation.

Reviewer comment (DocH_Monograph, page 9): The supplemental search of the non-English language databases is appropriate. However, what is the rationale for saying that they were used primarily to identify null or no-effect studies? Does that mean that if a study was identified that showed an association it was not abstracted? Please be a bit more clear on this.

Reviewer comment (DocJ_Monograph, page 7): In addition, can it be assumed that any non-English paper that met criteria, regardless of outcome, would have been included? While it is understandable why the confirmatory search was done, it could be perceived as biased to only search for and include papers with null findings.

explored before two were identified, 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 epidemiological 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 epidemiological 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 is available in the protocol (https://ntp.niehs.nih.gov/go/785076)]. Publications retrieved were compared with 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 2.

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. A primary goal of the screening of the newly retrieved human references in the supplemental search of Chinese databases was to identify studies that evaluated primary neurodevelopmental or cognitive outcomes (i.e., learning, memory, and intelligence) that may have been missed in previous searches that did not include the Chinese databases. A secondary goal was to examine whether the non-English-language studies on the Fluoride Action Network website (http://fluoridealert.org/)—a site used as another resource to identify potentially relevant studies because it is known to index fluoride publications—had been selectively presented to list only studies reporting effects of fluoride. Newly retrieved human references were reviewed to identify studies that may have been missed using previous approaches. Studies identified that evaluated primary neurodevelopmental or cognitive outcomes were included and either translated or reviewed by an epidemiologist fluent in Chinese.

Commented [A16]: This sentence reflects revisions to introduce the Fluoride Action Network and to clarify that the site was used as another resource because it is known to index fluoride publications, in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 6): This is the first time the "Flouride Action" website is mentioned (and the actual hyperlink appears in the subsequent section). It may be helpful to the reader to provide some rationale for why this website was specifically targeted.

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.

Unpublished Data

Although no unpublished data were included in the review, 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 (https://ntp.niehs.nih.gov/go/785076) for more details].

Study Selection

Evidence Selection Criteria

In order to be eligible for inclusion, studies had to satisfy eligibility criteria that reflect the PECO statements in Table 1, Table 2, and Table 3.

The following additional exclusion criteria were applied [see protocol (https://ntp.niehs.nih.gov/go/785076) for additional details]:

(1) Case studies and case reports. Although there are various definitions of 'case study' and 'case report,' the terms are used here to refer to publications designed to share health-related events on a single subject or patient with a disease, diagnosis, or specific outcome in the presence of a specific exposure (see Table 4 for study design definitions).

(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, theses, dissertations, and other non-peer-reviewed 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 (the 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 (Sciome) (Howard et al. 2020). While the human screeners review studies, SWIFT-Active Screener aids in this process by employing a machine-learning software program 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

Commented [A17]: Change made in response to the [redacted] comment below; see DocB2_Monograph for detailed response:

Reviewer comment (DocB2_Monograph, page 4): The URL was already noted above.

Commented [A18]: This sentence reflects revisions to state that no unpublished data were included in the monograph, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page6): Why would you exclude conference abstracts, theses, dissertations, and other non-peer reviewed reports, but include unpublished data?

Commented [A19]: This new sentence was added to clarify case studies and other study design definitions in response to the [redacted] Reviewer comment below. In addition, definitions of all observational study designs were added as footnotes to Table 4; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page5): As part of eligibility criteria, what is a "case study" and how does it differ from a case report? Did you have a minimum sample size? The best [redacted] could glean is >1. Table 4 implies that case series were included. Does case series mean a single group study (all had the same exposure) or a series of cases? In either case, how are these relevant? Please make explicit what the difference is between a cohort study and a cross-sectional study. I'm assuming you require cohort studies to be longitudinal, but this should be stated explicitly. Do you have a minimum duration of follow-up to count as a longitudinal study?
All of this can/should be addressed by adding Study Design rows to Tables 1-3.

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® (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.

Evaluation of SWIFT-Active Screener Results

During the initial title and abstract screening of 20,883 references using SWIFT-Active Screener, approximately 38%[8] of the studies were manually screened in duplicate to identify an estimated 98.6% of the predicted number of relevant studies using the software's statistical algorithm (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 were identified and moved to full-text review. The SWIFT-Active Screener 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, the evaluation team assessed the SWIFT-Active screening results to gain a better understanding of the relevance of the last group of studies that was screened before 98% predicted recall (i.e., 98% of the predicted number of relevant studies were identified). The goal was to determine the likelihood of having missed important studies by not screening all of the literature. To do this, the evaluation team examined subsets of studies screened in SWIFT-Active Screener 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, it was estimated that the use of SWIFT-Active Screener may have resulted in missing

[8]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.

one to two relevant human studies and one to two relevant animal studies with primary neurodevelopmental or cognitive outcomes. Therefore, the use of SWIFT-Active Screener saved considerable time and resources and is expected to miss very few potentially relevant publications.

Screening of the May 2020 Literature Search Update

For the May 1, 2020, literature search, only primary human epidemiological studies were identified for data extraction. 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 confidence ratings and 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. All included studies from the May 2020 literature search update appear in Appendix C; however, other than the primary human epidemiological studies, data from the new studies were not extracted unless they would materially advance the findings.

Note that NTP is aware of a conference abstract by Santa-Marina et al. on a Spanish cohort study that looked at fluoride exposure and neuropsychological development in children (Santa-Marina et al. 2019). The evaluation team conducted a targeted literature search in April 2021 to see whether the data from this study had been published. When no publication was found, the evaluation team contacted the study authors to inquire about the publication of their data. The response from the study authors indicated that the study report was being finalized but had not yet been sent to a journal for review; therefore, it was not considered here.[9]

Commented [A20]: This paragraph and footnote 9 reflect revisions to acknowledge the potential interest in Ibarluzea et al. (2021), in response to the [redacted] Reviewer comment below. Note the Ibarluzea study is addressed here and the Aggeborn study is addressed in the response to comments document; see DocA1_Monograph for detailed response:

Reviewer comment:
Issue: New evidence Recommendation
(DocA1_Monograph, page 7): The Ibarluzea and Aggeborn & Oehman studies should be evaluated and included when assessing the evidence, similar to the 15 additional studies from the Chinese databases. [redacted] also recommends NTP include a comparison between Ibarluzea et al., 2021, and Green et al., 2019, because both studies investigate fluoride exposures at levels used for water fluoridation.

Supplemental Chinese Database Searches and Human Epidemiological Studies

Supplemental searches were conducted in non-English-language databases (CNKI and Wanfang). Of the 910 references that were identified in the supplemental Chinese database searches, 13 relevant studies published in Chinese with primary neurobehavioral or cognitive outcomes were identified during title and abstract screening (which were not identified through the main literature searches). Full texts were not found for four studies after an extensive search. The remaining nine studies for which full texts were retrieved were included and were either professionally translated or evaluated by an epidemiologist fluent in Chinese for the data extraction and quality assessment steps described below. If necessary, author inquiries were conducted in Chinese to obtain missing information relevant to the assessment of the key risk-of-bias questions described below.

[9]NTP is aware that this study was published after April 2021 (Ibarluzea et al. 2021) and, therefore, is not included in this monograph because it is beyond the dates of the literature search. Even if it had been published earlier, the study would not have contributed to the body of evidence on children's IQ because the authors assessed other neurodevelopmental or cognitive effects, specifically the association between fluoride exposure and neuropsychological development in children aged 1 year using the Mental Development Index (MDI) of the Bayley Scales of Infant Development and in children aged 4 years using the General Cognitive Index (GCI) of the McCarthy Scales of Children's Abilities (MSCA). The study will be examined as part of the NTP meta-analysis, which is being prepared as a separate report for publication.

Data Extraction

Extraction Process

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

Data Availability

Data extraction was completed using the Health Assessment Workspace Collaborative (HAWC), an open-source and freely available web-based application.[10] 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 because they do not provide specific information on thyroid hormone levels that would inform whether a thyroid-mediated mechanism was involved in fluoride-associated changes in neurodevelopment. 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. Thyroid data were not extracted for animal studies due to inconsistency in the available data in humans. In vitro studies were evaluated, although data were not extracted from these studies as none of the findings were considered informative with respect to biological plausibility. The data extraction results for included studies are publicly available and can be downloaded in Excel format through HAWC (https://hawcproject.org/assessment/405/) (NTP 2019). 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 published 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 animal studies published prior to 2015 because these were reviewed in the NTP (2016) assessment.

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

Commented [A21]: This sentence reflects revisions to increase clarity regarding the exclusion of topics for full evaluation, in response to related comments from [redacted] Reviewers below; see DocG_Monograph and DocI_Monograph, respectively, for detailed responses. *Note: two additional sentences were revised toward the end of this paragraph to also address this issue as described in the next comment bubble.*

Reviewer comment (DocG_Monograph, page 23): A clearer statement is needed up front, ideally in the Introduction about what topics were covered by full systematic review (which are a small subset of topics of interest) and why. It's very confusing to read through repeated descriptions of topics which are not being reviewed. As an example, it's disconcerting to repeatedly see that thyroid function is an outcome of interest (without an explanation as to why this is of interest to a review of neurodevelopmental and cognitive health effects) and then to come across the statement (page 13) that "Thyroid data were … not extracted." It's difficult to pick out and follow the reasoning for excluding most topics from full evaluation. The timing of and reasoning for the decisions to focus the systematic review on just "high quality" pediatric studies is unclear.

Reviewer comment (DocI_Monograph, page 11): Page 13: there is a statement about studies 'evaluating only goiters or thyroid size were not extracted." If so, shouldn't they be p[a] …

Commented [A22]: The following two sentences reflect revisions to increase clarity regarding the exclusion of topics for full evaluation, and that in vitro studies were not summarized because it was considered unlikely that this literature would provide sufficient information to inform an action of fluoride on neurodevelopment, in response to comments from [redacted] Reviewers listed below; see DocG_Monograph and DocI_Monograph for detailed responses:

Reviewer comment (DocG_Monograph, page 23): A clearer statement is needed up front, ideally in the Introduction about what topics were covered by full systematic review (which are a small subset of topics of interest) and why. It's very confusing to read through repeated descriptions of topics which are not being reviewed. As an example, it's disconcerting to repeatedly see that thyroid function is an outcome of interest (without an explanation as to why this is of interest to a review of neurodevelopmental and cognitive health effects) and then …

Commented [A23]: This sentence reflects revisions to clarify that the statement applies to animal studies, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 7): Based on the Methods section, it appears that "high quality" pediatric studies from prior to 2015 would have been excluded in the current analyses. As written (e.g., on page 14 at the end of the Data Extraction section), it seems that older data were simply ignored (without justification). However, the Results (e.g., Figure 4) includes older studies.

Quality Assessment of Individual Studies

Risk of bias was assessed for individual studies using the OHAT risk-of-bias tool (https://ntp.niehs.nih.gov/go/riskbias) 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). When evaluating the risk of bias for an individual study, the direction and magnitude of association for any specific bias is considered.

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 prespecified 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 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 these issues are generally considered to have a greater impact on estimates of the effect size or on the credibility of study results in environmental health studies. There are three Key Questions for observational human studies: confounding, exposure characterization, and outcome assessment. Based on the complexity of the possible responses to these questions in epidemiological studies, considerations made and methods used for evaluating the Key Questions are provided below. There are 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.

Risk-of-bias Considerations for Human Studies

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) 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 have the greatest potential impact on the results. The other risk-of-bias questions, including selection of study participants, were also considered and were used to identify any other risk-of-bias concerns that may indicate serious issues with a study that could cause it to be considered high risk of bias. No study was excluded based on concerns for risk of bias; however, the low risk-of-bias studies generally drive the ratings on confidence in the results across the

Commented [A24]: This URL was added in response to the [redacted] Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 13): Has the OHAT been published? If so, it should be referenced. Since it's a critical tool in this review, it needs to be further described. What other QA tools are available and why weren't they used? Were the Cochrane Review recommendations for assessment of the risk of bias in research studies followed?

Commented [A25]: This sentence reflects revisions to clarify that in addition to the three key risk-of-bias questions, the impact of selection bias and other risk-of-bias questions were considered, in response to related comments from [redacted] Reviewers shown below. Also note the limitation in the *Discussion* section for high risk of bias studies "*Studies did not provide sufficient direct information (e.g., participation rates or methods for selection) to evaluate selection bias.*"; see DocH_Monograph and DocJ_Monograph, respectively, for detailed response.

Reviewer comment (DocH_Monograph, page 10): The focus on confounding, exposure characterization and outcome assessment are, as indicated, the key components of evaluating observational research. The other parameter is whether the participants represent the population from which they are recruited, i.e. selection bias. In prospective cohort studies this is not an issue, as the population is really the combination of those exposed and non-exposed. For cross sectional studies, this is a bit trickier, as the participants may reflect a select group within the overall population. For studies based on national or regional registries, such as the Canadian studies, this is less of a problem, but for others there is the possibility of bias, and the direction of such bias is difficult to predict. As [redacted] looked at the studies, the vast majority do not address this issue, but [redacted] believe that it is worth a discussion or at least a mention that the possibility of selection bias is real.

Reviewer comment DocJ_Monograph, page 14): In general, it is difficult to understand how cross-sectional studies that adjusted for few, or no, confounders, employed somewhat indirect measures of fluoride exposure (or did not fully capture all sources of exposures to fluoride), or had concerns related to selection bias, were designated as "low risk of bias." If, for example, some confounders were accounted for in the design or analysis, other than statistical adjustment, it may be worth noting that on Table 6 (otherwise, it appears that many papers accounted for no confounders).

body of evidence. Human evidence was evaluated with and without high risk-of-bias studies to assess the impact of these studies on confidence in the association.

High risk-of-bias studies: Studies rated probably high risk of bias for at least two key risk-of-bias questions or definitely high for any single question are considered studies with higher potential for bias (i.e., high risk-of-bias studies) and to be of low quality. Studies could also be considered high risk of bias if rated probably high risk of bias for one key risk-of-bias question along with other concerns, including potential for selection bias and concerns with statistical methods.

Low risk-of-bias studies: The remaining studies (i.e., other than the high risk-of-bias studies) were considered to have lower potential for bias (i.e., low risk of bias) and to be of high quality. Appendix E describes strengths and limitations of the low risk-of-bias/high-quality studies identified during the assessment and clarifies why they are considered to pose low risk of bias. Details on the statistical analyses are provided in the "Other potential threats" domain in order to evaluate the adequacy of the statistical approach for individual studies.

Given the number of non-English-language studies in this assessment, the potential for the translation to introduce bias was examined as described below, and it was determined that translation of non-English-language studies did not impact evaluation of risk of bias. Thirty-two of 100 studies included in the entire human body of evidence on neurodevelopmental and cognitive effects were initially published in a foreign language (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 the key risk-of-bias questions. Therefore, in order to assess whether 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 low risk-of-bias group of studies were reviewed by a team member ~~fluent in~~ with Chinese as first language to determine whether the translations were accurate and whether 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 of the translations on the key risk-of-bias concerns.

Confounding

Covariates were determined a priori based on factors that are associated with neurodevelopment or cognition and could be related to fluoride exposure. Covariates that were considered key for all studies, populations, and outcomes included age, sex, and socioeconomic status (e.g., maternal education, household income, marital status, crowding). Additional covariates 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., attention deficit hyperactivity disorder [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

Commented [A26]: The following sentence was added to clarify the definition of "high risk-of-bias studies". In addition, the detailed assessments of and justifications for risk-of-bias ratings for the key studies are provided in Appendix E (Details for Low Risk-of-bias Studies) to address the potential concern of confounding and exposure classification in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment: (DocJ_Monograph, page 18) Page 48-49, Assessment of Risk of Bias: While the studies noted as "low risk of bias" are certainly lower risk than the studies noted as "high risk of bias," it appears that the evidence base is still subject to a number of important risks, particularly related to confounding and exposure classification (i.e., are they "low risk" or "lower risk"?).

Commented [A27]: The following four sentences reflect revisions to clarify that all translated studies were originally published in Chinese and a review team member with Chinese as first language confirmed the accuracy of the translations, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page7): The description of the translated Chinese articles (page 15) needs to be written in the active voice to better describe who was confirming the accuracy of the translation and how. What about other languages (and what were they)?

Commented [A28]: This sentence reflects revisions to clarify that age and sex are important potential confounders regardless of life stage, in response to the [redacted] Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 8): The approach to assess risk of bias was clearly described. A brief discussion is needed about critical confounders, including a biological exposure measure for tobacco use or exposure, such as serum cotinine, and parental IQ for the child studies If there are unique confounders for child and adult studies, this needs to be articulated. It currently appears that there are no unique confounders for child and adult.

(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 the confounding domain, studies were not required to address every important covariate 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 covariates considered important for the specific study population and outcome. For example, studies of populations in China, India, and Mexico, where there is concern about co-exposures to high fluoride and high arsenic, were required to address arsenic. If the authors did not directly specify that arsenic exposures were evaluated, groundwater quality maps were evaluated (https://www.gapmaps.org/Home/Public) in order to identify areas of China, India, and Mexico where arsenic is a concern (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. ~~however, it should be noted that arsenic may be associated with neurodevelopmental effects at concentrations below 10 µg/L~~.

Commented [A29]: This sentence and text throughout the monograph reflect revisions to change the word 'confounder' to 'covariate', in response to several related comments from the same [redacted] Reviewer listed below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 5): RE: confounding and covariates. Recent thinking regarding confounding requires the use of directed acyclic graphs to define variables which are theoretically confounders (based on previous literature). Thus, some clarification is needed on how the set of three important confounders were selected, i.e. sex, child age and a measure of socioeconomic status. Indeed, based on literature from other potential neurotoxins (e.g. lead, polychlorinated biphenyls, phthalates) it seems ...

Commented [A30]: The removed text had initially been added in response to the [redacted] Reviewer comment below; however, as we were unable to appropriately support the statement with a reference, this statement has been removed. See DocH_Monograph for detailed response.

Reviewer comment (DocH_Monograph, page 10): For confounding, please see [redacted] remarks above. [redacted] do think that biological sex needs to be considered an effect modifier as in other studies of neurotoxins and neurodevelopmental outcomes. Further, as indicated later in the monograph, at ...

Exposure

Fluoride ion is rapidly absorbed from the gastrointestinal tract and is rapidly cleared from serum by distribution into calcified tissues and urinary excretion (IPCS 2002). There is general consensus that the best measures of long-term fluoride exposure are bone and/or tooth measurements, and other than measures of dental fluorosis, these were not performed in any of the studies reviewed in this document. Prolonged residence in an area with a given fluoride content in drinking water has been considered in many studies as a proxy for long-term exposure.

Commented [A31]: This paragraph was added in response to two related comments on toxicokinetics from the same [redacted] Reviewer listed below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 7): In general the report is comprehensive and includes all necessary material. Hence, [redacted] have no major additions or deletions. [redacted] have one small addition, which would be a discussion of the toxicokinetics of fluoride – this is necessary because the half life is relatively short, and a spot measure ...

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 of gestation), serum, individual drinking water, intake from infant formula, estimated total exposure dose, municipal drinking water (with residence information), evidence of dental or skeletal fluorosis, 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.

Commented [A32]: This sentence was added in response to the [redacted] Reviewer comments on serum fluoride below; see DocI_Monograph for detailed responses:

Reviewer comment (DocI_Monograph, page 7): A brief discussion of serum fluoride needs to be included – similar to the urinary fluoride description (page 16).

Reviewer comment (DocI_Monograph, page 12): Add a brief section on serum fluoride levels. Urinary fluoride levels is fully described, but serum has been omitted.

Urinary fluoride levels measured during pregnancy and in children include all ingested fluoride and are considered a valid measure to estimate total fluoride exposure (Villa et al. 2010; Watanabe et al. 1995); however, the type and timing of urinary sample collection are important to consider. 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 with 24-hour urine samples, spot urine samples are more prone to the influence of timing of exposure and can also be affected by differences in dilution; however, many studies attempted to account for dilution either by 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 (e.g.,

Commented [A33]: This sentence was added to address the best measures for assessing long-term fluoride exposure, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 16): Because the urine and serum biomarkers of fluoride represent relatively recent exposure, it is difficult to infer that the associations are from cumulative exposure without laying out the assumptions, i.e. long term residential history, similar habits of toothpaste use, etc.

accounting for dilution), studies that used this metric were generally considered to have probably low risk of bias for exposure.

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 urinary fluoride levels from some individuals to verify that there were differences in the fluoride exposure between groups. Studies that provided results by area and 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. Ideally, these studies would still need to consider and adjust for area-level clustering; however, these concerns are captured in evaluations of other potential threats to internal validity.

Commented [A34]: This sentence was added in response to the [redacted] Reviewer comment on clustering below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 11): A further concern with exposure assessment brought up in the previous review concerns the issue of clustering with regard to exposure. The authors of the monograph do a very nice job of addressing this issue as it was raised in the prior review, but pointing to the sensitivity analyses. [redacted] only concern remaining is that this is mentioned up front when the exposure characterization is discussed in the methods.

Outcome

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 many of the tests to measure neurodevelopment and cognitive function 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. If cross-sectional studies collected biomarker measurements at the time of an IQ assessment, this was considered indirect evidence that the outcome assessor would not have knowledge of the fluoride exposure unless there was also potential for the outcome assessor to have knowledge of varying levels of fluoride by study area. In cases wherein the study did not specify that the outcome assessors were blind, the study authors were contacted and asked whether the outcome assessors were, in fact, blind to exposure. When authors responded and indicated that outcome assessors were blind to exposure or that it was not likely that they would have had knowledge of exposure, this was considered direct or indirect evidence, respectively, that blinding was not a concern for those 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 4. OHAT Risk-of-bias Questions and Applicability by Study Design

Risk-of-bias Questions	Experimental Animal[a]	Human Controlled Trials[b]	Cohort[c]	Case-control[d]	Cross-sectional[e,]	Case Report/Case Series[f]
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

[a]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.
[b]Human Controlled Trials are studies in humans with controlled exposure (e.g., randomized controlled trials, non-randomized experimental studies).
[c]Cohort studies are observational studies in humans that examine a cohort prospectively or retrospectively over time. Although cohort studies may include longitudinal analyses, it is not a prerequisite of the cohort study design.
[d]Case-control studies are observational studies in humans that compare exposures of individuals who have a specific health effect or disease with exposures of controls who do not have the effect or disease. Controls generally come from the same population from which the cases were derived.
[e]Cross-sectional studies are observational studies in humans that examine the relationship between exposures and outcomes or health effects assessed contemporaneously. ~~e~~Cross-sectional studies include population surveys with individual data (e.g., NHANES) and surveys with aggregate data (i.e., ecological studies).
[f]A case report (or case study) is a descriptive study of a single individual or small group in which the study of an association between an observed effect and a specific environmental exposure is based on clinical evaluations and histories of the individual(s). A case series study in environmental epidemiology is designed to share health-related events on a collection of case reports on subjects with the same or similar health outcome(s) and environmental exposure(s).

Commented [A35]: Footnote c was revised and new footnotes d, e, and f were added to provide definitions of all observational study designs, in response to the [redacted] Reviewer comments below; see DocG_Monograph and DocJ_Monograph, respectively, for detailed response:

Reviewer comment (DocG_Monograph, page 5): As part of eligibility criteria, what is a "case study" and how does it differ from a case report? Did you have a minimum sample size? The best [redacted] could glean is >1. Table 4 implies that case series were included. Does case series mean a single group study (all had the same exposure) or a series of cases? In either case, how are these relevant? Please make explicit what the difference is between a cohort study and a cross-sectional study. [redacted] assuming you require cohort studies to be longitudinal, but this should be stated explicitly. Do you have a minimum duration of follow-up to count as a longitudinal study?
All of this can/should be addressed by adding Study Design rows to Tables 1-3.

Reviewer comment (DocJ_Monograph, page12): In general, it would be helpful to the reader to describe the included study designs earlier in the respective results sections (i.e., along with the total #s of included articles, describing the # of cross-sectional, prospective cohort, etc is important).

Commented [A36]: Footnote c on cohort studies was revised in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 14): It is unclear whether you are using the term prospective (cohort study) to mean prospective (as opposed to retrospective, it's correct meaning) or longitudinal (as opposed to cross-sectional). Please use the correct term. In any case, the reader needs to know both whether studies were prospective or retrospective and whether studies were longitudinal or cross-sectional.

Answers to the risk-of-bias questions result in one of the following four risk-of-bias ratings:

Table 5. The Four Risk-of-bias Rating Options

Symbol	Description
++	***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 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

This evaluation provides only a narrative review of the data; however, heterogeneity within the available evidence was evaluated to determine whether a quantitative synthesis (i.e., meta-analysis) would be appropriate. 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) reported a significant non-linear dose-response relationship between fluoride dose and intelligence with the relationship stated as most evident with exposures from drinking water above 4 mg/L (or 4 ppm) fluoride. 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). The meta-analysis conducted in

Commented [A37]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, page 13): Would revise because it is being submitted for publication and also [redacted] would not link it to further informing this since there are still questions about the meta-analysis.

association with this systematic review further informs this issue and will be ~~published seperately~~refined in preparation for a separate publication.

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 protocol (https://ntp.niehs.nih.gov/go/785076). 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). Potential upgrading of the confidence rating was considered for factors that increase confidence in the results (see column 3 of Figure 1). Short descriptions of the factors that can decrease or increase confidence in the body of evidence for human studies are provided below [see protocol (https://ntp.niehs.nih.gov/go/785076) for additional details related to the human body of evidence, as well as considerations for experimental animal studies].

Factors to Consider for Potential Downgrading

- Risk of bias: Addresses whether the body of evidence did not account for critical factors in study quality or design, including confounding bias, selection bias, exposure assessment, and outcome assessment. Consideration for downgrading the confidence rating is based on the entire body of evidence, and the evidence is downgraded when there is substantial bias across most studies that could lead to decreased confidence in the results and when the studies without substantial bias could not support the confidence rating. Individual studies are evaluated for risk of bias based on a set of criteria (as discussed above); magnitude and direction of the bias are also considered.
- Unexplained inconsistency: Addresses inconsistencies in results across studies of similar populations and design that can be determined by assessing similarity of point estimates and extent of overlap between confidence intervals or more formally through statistical tests of heterogeneity. Sensitivity analysis can be used to assess the impact of specific variables on the outcome. Inconsistencies that can be plausibly explained by characteristics of the studies (e.g., sex-associated differences) are typically not used to support a downgrade. A downgrade would only be applied when there is an inconsistency that cannot be explained and results in reduced confidence in the body of evidence.
- Indirectness: Addresses generalizability and relevance to the objective of the assessment. As outlined in the Objective and consistent with the population specified in the PECO statement, this systematic review evaluated the extent and quality of the evidence linking fluoride exposure to neurodevelopmental and cognitive effects in humans without restriction as to age, sex, geographic location, or life stage at exposure or outcome assessment. Furthermore, the review did not exclude subjects exposed in occupational settings. All exposure levels and scenarios encountered in

human studies are considered direct (i.e., applicable, generalizable, and relevant to address the objective of the assessment); therefore, a downgrade for indirectness would not be applied to bodies of evidence from human studies.

- Imprecision: Addresses confidence associated with variability in quantitative measures such as effect sizes. Typically, 95% confidence intervals are used as the primary method to assess imprecision, but considerations can also be made on whether studies were adequately powered. Meta-analyses can also be used to determine whether the data are imprecise. When a meta-analysis is not appropriate or feasible, imprecision can be based on variability around the effect estimate. A downgrade would occur if the body of evidence was considered to be imprecise based on a meta-analysis, or if serious or very serious imprecision was consistently present in the body of evidence. A downgrade is especially likely if imprecision raised questions as to whether an overall effect was significant.
- Publication bias: Addresses evidence of biased publication practices. Downgrade if one strongly detects publication bias. Publication bias is difficult to detect but may be evident if major sections of the research community are not publishing (e.g., absence of industry, academic, or government studies) on a topic or if there are multiple instances wherein data from conference abstracts are never published in peer-reviewed journals. In addition, there are methods included in conducting a meta-analysis to detect whether there is potential for publication bias, including the use of fit-and-trim models, which help identify how publication bias may affect the results of the meta-analysis. Although a meta-analysis is not included in this systematic review, there are two published meta-analyses (Choi et al. 2012; Duan et al. 2018) in addition to the one associated with this systematic review (manuscript in progress) that can be used to address publication bias.

Factors to Consider for Potential Upgrading

- Large magnitude of effect: Factors to consider include the outcome being measured and the dose or exposure range assessed. The confidence can be upgraded if the body of evidence is suggestive of a large magnitude of effect. GRADE provides guidance on what can be considered a large magnitude of effect based on relative risk (i.e., suggests one upgrade in confidence if relative risk is greater than 2 and two upgrades in confidence if greater than 5). However, not all studies provide data as a risk estimate, and smaller changes, such as increases in blood pressure, may have greater impact on health at the population level. Consideration for an upgrade is not based on a single study, and what constitutes a large magnitude of effect will depend on the outcome and the potential public health impact.
- Dose response: Patterns of dose response are evaluated within and across studies. Confidence in the body of evidence can be increased when there is sufficient evidence of a dose-response pattern across multiple studies.
- Consistency: Does not apply in this evaluation. The consideration of a potential upgrade for consistency is primarily for non-human animal evidence in which it would be applied to address increased confidence based on an observation of consistent effects across multiple non-human animal species. For human evidence, this factor would generally not be applied. Human studies are instead evaluated for

issues of consistency that could result in downgrading confidence for unexplained inconsistency (see "Factors to Consider for Potential Downgrading" above).

- Consideration of residual confounding: Applies to observational studies and refers to consideration of unmeasured determinants that are likely to be distributed unevenly across groups. Residual confounding can push results in either direction, but confidence in the results is increased when the body of evidence is biased by factors that counter the observed effect and would cause an underestimation of the effect. Confounding that would cause an overestimation of the effect is considered under the risk-of-bias considerations for decreasing confidence.

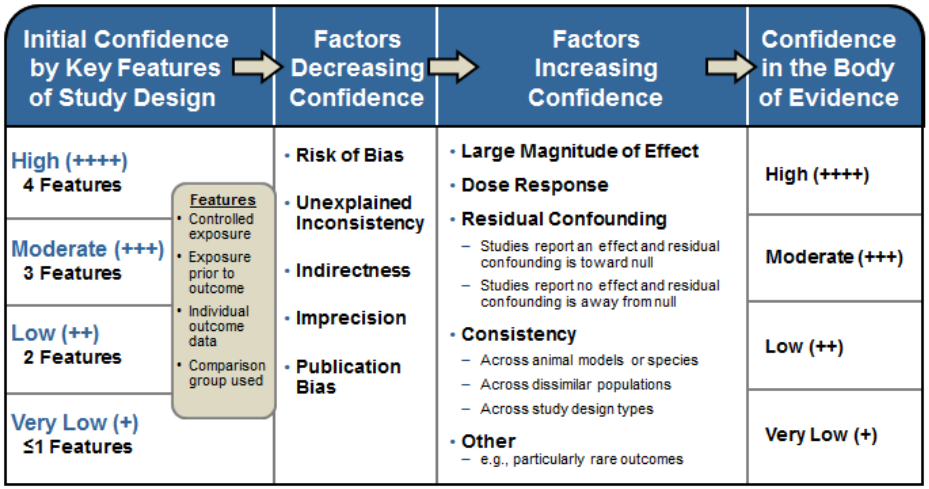


Figure 1. Assessing Confidence in the Body of Evidence

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.

Results

Literature Search Results

The electronic database searches retrieved 25,450 unique references with 11 additional references[11] identified by technical advisors or obtained by manually searching the Fluoride Action Network website or reviewing reference lists of published reviews and other included studies. During title and abstract screening, 1,036 references were moved to full-text review and 24,425 were excluded (11,402 by manual screening for not satisfying the PECO criteria and 13,023 based on the SWIFT-Active Screener algorithm). Among the 1,036 references that underwent full-text review, 547 studies were considered PECO-relevant (see Appendix C for list of included studies). A few studies assessed data for more than one evidence stream (human, non-human mammal, and/or in vitro), and several studies assessed more than one type of outcome (e.g., primary and secondary outcomes). Included studies break down as follows:

- 167 human studies (84 primary only; 13 secondary only; 5 primary and secondary; 8 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).

Additional details on the screening results are provided in Appendix C. These screening results are outlined in a study selection diagram that reports numbers of studies excluded at each stage and documents the reason for exclusion at the full-text review stage (see Figure 2) [using reporting practices outlined in ~~Moher et al. (2009)~~ Page et al. (2021)].

Commented [A38]: This footnote was added to further describe why and how studies were identified from other sources in response to related comments from [redacted] Reviewers listed below; see DocG_Monograph and DocJ_Monograph, respectively, for detailed responses:

Reviewer comment (DocG_Monograph, page 9): Page 22, can you provide a brief explanation for why the 15 additional identified references were missed by your literature searches?

Reviewer comment (DocJ_Monograph, page 11): Identifying 15 references through other sources seems somewhat high. Was there a need to adjust the original search strategy to capture those references?

Commented [A39]: Change made in response to the [redacted] comment below; see DocB1_Monograph for detailed response:

Reviewer comment (DocB1_Monograph, page 5): The authors are using an old version of the PRISMA flow diagram - The 2020 PRISMA flow diagram can be found here: https://prisma-statement.org/prismastatement/flowdiagram.aspx

[11]These 11 studies (9 human and 2 animal studies) were not identified through the electronic database searches, as they were not indexed in any of the electronic databases searched. Note that the supplemental search of non-English-language databases was designed in part to identify non-English-language studies that are not indexed in traditional bibliographic databases such as PubMed. It was successful in this goal, as multiple studies that were initially only identified through "other sources" were subsequently captured in the supplemental Chinese database search, leaving only 11 as identified through other sources.

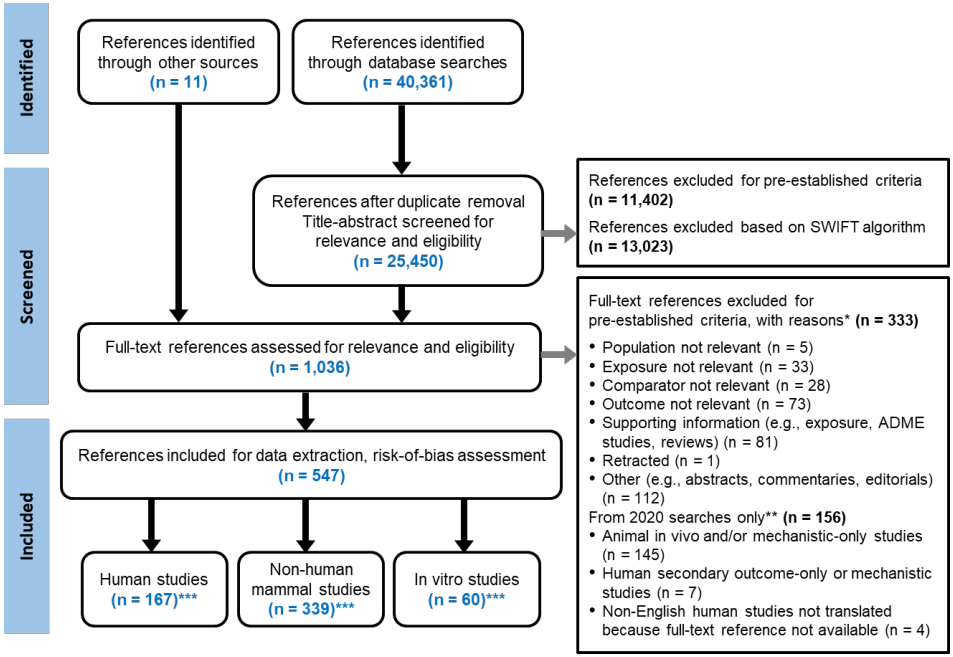


Figure 2. Study Selection Diagram[a]

[a]An interactive reference flow diagram is available here: https://hawcproject.org/summary/visual/assessment/405/Figure-2/.
*~~Includes s~~Studies from all literature searches conducted during the review excluded at the full-text level for pre-established criteria; see the Methods section for extraction and search update information. Studies may have been excluded for more than one reason; the first reason identified was recorded.
**~~Includes all studies from all~~ Studies excluded from the 2020 literature searches ~~not otherwise excluded~~ for reasons other than pre-established criteria; see the Methods section for extraction and search update information.
***Publications may contain more than one evidence stream, so the numbers will not total the 547 included studies.

Human Neurodevelopmental and Cognitive Data

The body of literature that evaluates the association between fluoride exposure and neurodevelopmental and cognitive effects in humans is relatively robust with a large number of studies (n = 100) that cover a wide array of endpoints (see Figure 3). Seventy-two human studies investigated IQ in children. Additional studies evaluated learning and memory (n = 9 studies) or other cognitive developmental effects (e.g., total neurobehavioral scores and total mental capacity index in children, cognitive impairment in adults; n = 15 studies).[12] For this review, the evidence in children and adults was evaluated separately to address potential differences in the health impact of fluoride exposure during development versus adulthood.

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

Commented [A40]: Figure 2 has been revised for accuracy, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 10): The numbers of abstracts in Figure 2 do not align with the text.

Commented [A41]: Footnotes * and ** were revised to provide additional clarity on the 333 excluded studies. Although the reviewer was not correct in their suggestion it pointed out the need for clarification in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 11): The top line of the excluded box should state that the 333 were from the original (pre-2020) search.

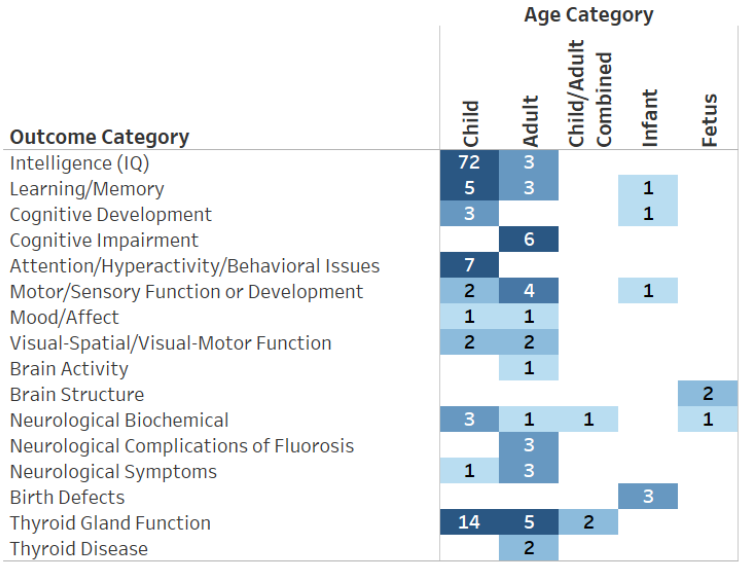


Figure 3. Number of Epidemiological Studies by Outcome and Age Categories[a]

[a]Interactive figure and additional study details are available at https://public.tableau.com/app/profile/ntp.visuals/viz/FluorideTableauDashboards/ReadMe.
Choi et al. (2015) used subtests of the omnibus IQ test reported by the authors as Wechsler Intelligence Scale for Children-Revised (WISC-IV) to evaluate visuospatial abilities (using block design) and executive function (using digit span). These endpoints are included in the intelligence (IQ) outcome category as they are subsets of the IQ tests.
Three additional publications based on subsamples (i.e., 50–60 children) of the larger Yu et al. (2018) cohort were identified (Zhao et al. 2019; Zhao et al. 2020; Zhou et al. 2019) and are not included in the counts of this figure.

Because the majority of studies evaluated intelligence, the following section focuses on IQ effects in children followed by separate discussions on other measures of cognitive function and neurobehavioral effects in children and cognitive effects in adults. Studies that evaluated mechanistic data in humans, including effects on the thyroid, are discussed in the Mechanistic Data in Humans section. Note that a few studies were identified on congenital neurological malformations and neurological complications of fluorosis; however, they are not considered further due to the limited number of studies and the heterogeneity of outcomes evaluated in those studies.

Commented [A42]: This sentence reflects revisions that a few studies on these other health outcomes were identified, which is more in line with other statements in the *Results* section, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 11): On page 24, [redacted] have trouble with the Results statement "Congenital neurological malformations and neurological complications of fluorosis are not considered further due to the limited number of studies and the heterogeneity of outcomes evaluated in these studies." This belongs in the Methods, complete with a full explanation for criteria used to or not to report/consider.

IQ in Children

Seventy-two epidemiological studies were identified that evaluated the association between fluoride exposure and children's IQ. Nineteen of the 72 IQ studies were determined to have low potential for bias (i.e., were of high quality). Looking across the literature, there has been a progression over the years in the quality of studies conducted to assess the association between fluoride exposure and IQ in children, with more recent studies including better study designs, larger sample sizes, and more sophisticated statistical analysis. Older studies often had limitations related to study design or methods, and most of the high risk-of-bias studies (i.e.,

studies of low quality) were published prior to the 2006 NRC evaluation of fluoride in drinking water. In contrast, 18 of the low risk-of-bias studies were published after the 2006 NRC evaluation of fluoride in drinking water, and over half of those were published between 2015 and 2020 (Figure 4).

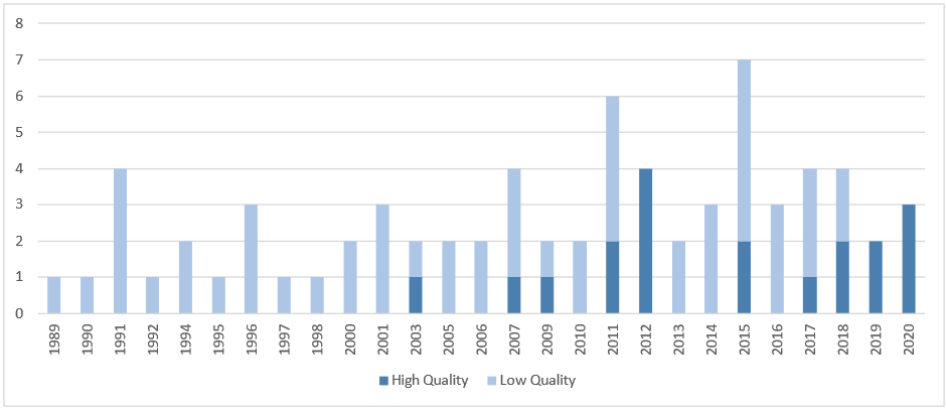


Figure 4. Number of High- and Low-quality Studies of Fluoride Exposure and IQ in Children by Year of Publication

Several characteristics of recent studies contribute to higher study quality in the overall body of literature on children's IQ and fluoride, including:

- Demonstration that exposure occurred prior to outcome assessment (an important factor when considering confidence in study results; see Figure 1) either by study design (e.g., for prospective cohort studies) or analysis (e.g., prevalence of dental fluorosis in children, limiting study populations to children who lived in the same area for long periods of time).
- Improved reporting of key study details that are necessary to evaluate study quality and allow for a more precise analysis of risk of bias.
- Increased consideration of key covariates (e.g., socioeconomic status) including potential co-exposures (e.g., arsenic or lead intake).
- Increased use of individual-level exposure measures (urine or water) as well as prenatal fluoride exposure to assess either individual-level fluoride exposure or—if still using group-level data—to confirm that regions being compared had differences in fluoride exposure.
- Utilization of more sophisticated sampling techniques for the study populations (e.g., stratified multistage random sampling).
- Application of more sophisticated regression approaches (e.g., piecewise linear regression models, multi-level regression with random effects, or generalized additive models for longitudinal measurements of fluoride).

- For studies using individual-level exposure measures, application of more sophisticated regression techniques to account for clustering at the cohort level by using cohort as a fixed or random effect and by accounting for numerous covariates that capture the cohort effect.

In addition, newer studies represent more diverse study populations across several countries (Figure 5), whereas all identified peer-reviewed studies that were published prior to 2006 took place in a single country (China). The majority of high-quality, low risk-of-bias studies exhibit these important study design and analysis characteristics, as discussed further in subsequent sections.

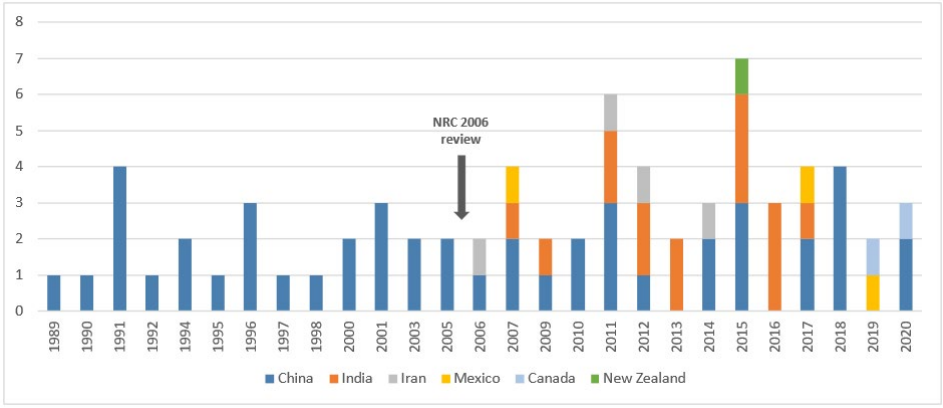


Figure 5. Number of Studies of Fluoride Exposure and IQ in Children by Country and Year of Publication

All available studies were considered in this evaluation; however, review of the body of evidence focused on the high-quality, low risk-of-bias studies for two main reasons. First, there are fewer limitations and greater confidence in the results of the high-quality studies. Second, there are a relatively large number of high-quality studies (n = 19), such that the body of evidence from these studies could be used to evaluate confidence in the association between fluoride exposure and changes in children's IQ. Therefore, the remainder of the discussion on IQ in children focuses on the 19 studies with low risk of bias. The high risk-of-bias studies are discussed briefly relative to their overall support of findings from the low risk-of-bias studies.

Low Risk-of-bias IQ Studies

Overview of Studies

Nineteen studies (3 ~~longitudinal~~ prospective cohort and 16 cross-sectional studies) with low potential for bias evaluated the association between fluoride exposure and IQ in children (see Quality Assessment of Individual Studies section for methods on determining which studies pose low risk of bias). These IQ studies were conducted in 15 study populations across 5 countries

Commented [A43]: This sentence reflects revisions to include information on counts of studies per study design, in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 12): In general, it would be helpful to the reader to describe the included study designs earlier in the respective results sections (i.e., along with the total #s of included articles, describing the # of cross-sectional, prospective cohort, etc is important).

Commented [A44]: The word "longitudinal" was previously added here and two other places in the monograph in response to the [redacted] Reviewer comment below; however, upon further consideration, we decided it would be more clear and consistent to add definitions of all observational study designs as footnotes to Table 4. Therefore, the new footnotes now clarify that cohort studies are observational studies in humans that examine a cohort prospectively or retrospectively over time, and although cohort studies may include longitudinal analyses, it is not a prerequisite of the cohort study design. See DocG_Monograph for detailed response.

Reviewer comment (DocG_Monograph, page 14): It is unclear whether you are using the term prospective (cohort study) to mean prospective (as opposed to retrospective, it's correct meaning) or longitudinal (as opposed to cross-sectional). Please use the correct term. In any case, the reader needs to know both whether studies were prospective or retrospective and whether studies were longitudinal or cross-sectional.

Commented [A45]: This sentence reflects revisions to refer the reader back to the *Methods* section that describes the risk-of-bias assessment for human studies, in response to the [redacted] Reviewer comment below; see DocK_Monograph for detailed response:

Reviewer comment (DocK_Monograph, page 4): It might be useful to have reminder, or reference back to the section in the text where the risk of bias information for human and animal studies is described in the methods (page 18), prior to presentation of the low risk of bias results for humans (page 28) and animals (page 67).

and included more than 7,000 children. Specifically, of the 19 low risk-of-bias studies of IQ in children:

- ten were conducted in four areas of China on seven study populations,[13]
- three were conducted in three areas of Mexico on three study populations,
- two were conducted in Canada using the same study population,
- three were conducted in three areas of India on three study populations, and
- one was conducted in Iran.

Most studies measured fluoride in drinking water (n = 15) and/or urine (child or maternal) (n = 15). Two studies measured fluoride in serum. The IQ studies used a variety of tests to measure IQ. Because IQ tests should be culturally relevant, the tests used often differed between studies, reflecting adjustments for the range in populations studied (e.g., western vs. Asian populations). In some cases, different IQ tests were used to study similar populations. Overall, these studies used IQ tests that were population- and age-appropriate.

Table 6 provides a summary of study characteristics and key IQ and fluoride findings for the 19 low risk-of-bias studies (organized by country and then by year). Several of these studies conducted multiple analyses and reported results on multiple endpoints. The purpose of the table is to summarize key findings (independent of whether an association is indicated) from each study and is not meant to be a comprehensive summary of all results from each study. For each study, results are summarized for each exposure measure assessed, but results from multiple analyses using the same exposure measure may not be presented for all studies unless multiple analyses yielded conflicting results. See Appendix E for additional information on each study in Table 6, including strengths and limitations, clarifications for why studies are considered to pose low risk of bias, and information regarding statistical analyses, important covariates, exposure assessment, and outcome assessment.

[13]In this document, "study population" refers to a defined population on which an original body of research was conducted. The published work drawn from that original body of research is often referred to as a "study." IQ studies that report on the same study populations are identified in Table 6.

Commented [A46]: This sentence was revised to clarify that Table 6 was organized by country and then by year in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 12): In general, studies designated as "low risk of bias" were interpreted correctly. I have a few suggestions as to how to clarify many of the points made.
While the results are generally consistent (table 6) it would be useful to present the results based on the exposure metric used. For example, studies using fluoride concentrations in "high" and "low" areas could be grouped together to illustrate the change in IQ points. Additionally, the actual IQ test used could also be used to group studies within exposure metric. There are clear differences in the scoring for the Raven and the WASI/WPPSI, for example and these are hard to tell from the presentation.

Commented [A47]: This sentence, and several sentences throughout the monograph, reflect revisions to replace 'effect' with 'association', in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 16): Also note that observational studies (with rare exceptions) do not provide evidence of an "effect", only of an association. Please use the term judiciously or not at all.

Commented [A48]: This sentence was added to direct readers to Appendix E for additional information such as statistical methods not provided in Table 6 in response to the [redacted] Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 17): Table 6 could include the following: 1) statistical methods; 2) confounders, particularly exposure to other known neurotoxicants, and how they were measured; 3) might rename 'Assessment timing' to age of participants or just combine the information with the location/subject's column

Table 6. Studies on IQ in Children[a]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
China					
Xiang et al. (2003a)[d]	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 vs. endemic fluorosis)	Children (ages 8–13 years)	IQ: Combined Raven's Test for Rural China	Significant dose-related association of fluoride on IQ score based on drinking water quintile levels with significantly lower IQ scores observed at water fluoride levels of 1.53 mg/L or higher; % of subjects with IQ <80 was significantly increased at water levels 2.46 mg/L or higher; significant inverse correlation between IQ and urinary fluoride (Pearson correlation coefficient of −0.164); mean IQ scores for children in non-endemic region (100.41 ± 13.21) significantly higher than endemic region (92.02 ± 13.00) No statistical adjustment for covariates
Ding et al. (2011)	Cross-sectional Inner Mongolia (Hulunbuir City)/elementary school children [331]	Children's urine Range: 0.1–3.55 mg/L Drinking water (reported but not used in analyses) Mean (SD): 1.31 (1.05) 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 decrease in IQ score of 0.59 points; 95% CI: −1.09, −0.08) Adjusted for age
Xiang et al. (2011)[d]	Cross-sectional Wamiao and Xinhuai villages (Sihong County)/school children [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 linear trend across 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 associations at ≥0.05 mg/L serum fluoride Adjusted for age and sex

Commented [A49]: This text and text throughout the monograph reflect revisions to change the term 'gender' to 'sex', in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 4): Use of the term "gender" to denote sex differences is not in line with current usage. Gender refers to the socially constructed variable, while sex refers to the biological variable. Please change.

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
Wang et al. (2012)[d]	Cross-sectional Wamiao and Xinhuai villages (Sihong County)/school children [526]	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 vs. endemic fluorosis) Drinking water (reported for villages but not used in analyses) Mean (SD): 0.36 (0.11) (control), 2.45 (0.80) (high fluoride) mg/L	Children (ages 8–13 years)	IQ: Combined Raven's Test for Rural China	Significantly lower mean IQ in the endemic versus non-endemic regions, as reported in Xiang et al. (2003a); when high-exposure group was broken into four exposure groups based on fluoride intake, a dose-dependent decrease in IQ and increase in % with low IQ observed; significant correlation between total fluoride intake and IQ ($r = -0.332$); for IQ <80, adjusted OR of total fluoride intake per 1-mg/(person/day) was 1.106 (95% CI: 1.052, 1.163) Adjusted for age and sex
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 (block design and digit span)	Compared to normal/questionable fluorosis, presence of moderate/severe fluorosis significantly associated with lower total (adjusted $\beta = -4.28$; 95% CI: −8.22, −0.33) and backward (adjusted $\beta = -2.13$; 95% CI: −4.24, −0.02) digit span scores; linear associations between total digit span and log-transformed urinary fluoride (adjusted $\beta = -1.67$; 95% CI: −5.46, 2.12) and log-transformed drinking water fluoride (adjusted $\beta = -1.39$; 95% CI: −6.76, 3.98) observed but not significant; forward digit span had similar results as backward and total but was not statistically significant; block design (square root transformed) not significantly associated with any measure of fluoride exposure Adjusted for age and sex, parity, illness before 3 years old, household income last year, and caretaker's age and education

Commented [A50]: This text and text throughout the monograph reflect revisions to update units of change in effect estimate per change in fluoride exposure or add cutoffs for categorical outcomes, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 4): The estimated regression coefficients from the studies need to be presented more clearly. For example, many times there is no reference, e.g. increase (or decrease) in score per 1 mg/L F in urine. Further, for the presentation of odds ratios, it is not clear what the dichotomous (or categorical) outcome variable is (e.g. IQ below 50). These suggestions are for clarity as well as for correctness.

Commented [A51]: This summary of the study results, and text throughout the monograph reflect revisions to change the term 'correlation' to 'association' when a regression coefficient was used, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 4): In general the scientific information presented, including the data in tables and figures, is technically correct and clearly and objectively presented. Specific comments regarding the general evaluation of studies:
The use of the term "correlation" is confusing ([redacted] have marked this several times in the document and tables, but there are also other occurrences). Correlation is generally used to denote a correlation coefficient (either Pearson or Spearman); however, [redacted] believe it has been used to denote the estimated regression coefficients (more on this below). [redacted] would recommend changing the terminology for clarity.

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
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 children's serum fluoride ($r = -0.47$) and urinary fluoride ($r = -0.45$); significant difference in mean IQ score for high-fluoride area (defined as >1 mg/L in drinking water; 102.33 ± 13.46) compared with control area (109.42 ± 13.30); % of subjects with IQ <90 significantly increased in high-fluoride area (28.7%) vs. low-fluoride area (8.33%); not significantly correlated with water fluoride Adjusted for age and sex, if applicable
Cui et al. (2018)	Cross-sectional Tianjin City (districts Jinghai and Dagang)/school children [323]	Children's urine Median (Q1–Q3): 1.3 (0.9–1.7) mg/L (boys), 1.2 (0.9–1.6) mg/L (girls)	Children (ages 7–12 years)	IQ: Combined Raven's Test for Rural China	Significant association between IQ score and log-transformed urinary fluoride (adjusted $\beta = -2.47$; 95% CI: −4.93, −0.01) Adjusted for age, mother's education, family member smoking, stress, and anger

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
Yu et al. (2018)[c,f]	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 areas (>1.0 mg/L; 106.4 ± 12.3 IQ) compared to the normal water fluoride areas (≤1.0 mg/L; 107.4 ± 13.0); distribution of the IQ scores also significantly different (p = 0.003); every 0.5-mg/L increase in water fluoride was associated with a decrease of 4.29 in IQ score (95% CI: −8.09, −0.48) when exposure was between 3.40 and 3.90 mg/L; no significant association between 0.2 and 3.40 mg/L; every 0.5-mg/L increase in urinary fluoride was associated with a decrease of 2.67 in IQ score (95% CI: −4.67, −0.68) between 1.60 and 2.50 mg/L but not at levels of 0.01–1.60 mg/L or 2.50–5.54 mg/L. Adjusted for age and sex, maternal education, paternal education, and low birth weight
Cui et al. (2020)	Cross-sectional Tianjin City (all districts)/school children (potentially some overlap with Cui et al. (2018)) [498]	Children's urine <1.6–≥2.5 mg/L	Children (ages 7–12 years)	IQ: Combined Raven's Test	Decreasing mean (± SD) IQ score with increasing urinary fluoride levels (statistical significance not reached 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 covariates

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
Wang et al. (2020b)[c]	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 urinary fluoride concentrations in boys and girls combined based on both quartiles and continuous measures (water: 1.587 decrease in IQ score per 1-mg/L increase; urine: 1.214 decrease in IQ score per 1-mg/L increase); no significant effect modification of sex Adjusted for age and sex, BMI, maternal education, paternal education, household income, and low birth weight
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 log-transformed fluoride and IQ scores (full-scale IQ adjusted βs of −10.2 [water] and −16.9 [urine]; CIs not reported); arsenic also present, but the association with arsenic was smaller (full-scale IQ adjusted βs of −6.15 [water] and −5.72 [urine]; CIs not reported) Adjusted for blood lead, mother's education, SES, height-for-age z-scores, and transferrin saturation

Commented [A52]: This summary of the study results reflects revisions to make clear that the association between arsenic exposure and children's IQ was smaller in magnitude than the association between fluoride exposure and children's IQ, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 13): At times, associations are presented as different when other covariates are controlled. [redacted] presume that these assessments were made by inspection of the results in the studies, but should either be backed up with statistical testing or admitted that they were made by inspection. For example, in table 6 the study by Rocha-Amador, et al states that the estimated associations between fluoride and the full scale IQ (WISC) were smaller when arsenic was controlled, the estimated betas are given, but there is no indication whether the differences are statistically different.

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
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	Significantly lower child IQ score per 0.5-mg/L increase in maternal urinary fluoride (adjusted β = −2.50; 95% CI: −4.12, −0.59); no significant association with children's urine Adjusted for sex, gestational age; weight at birth; 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
Soto-Barreras et al. (2019)	Cross-sectional Chihuahua/school children [161]	Children's urine Range: 0.11–2.10 mg/L Drinking water Range: 0.05–2.93 mg/L Fluoride exposure dose (summary statistics not reported) Fluorosis index (summary statistics not reported)	Children (ages 9–10 years)	IQ: Raven's Colored Progressive Matrices	No significant difference in urinary fluoride, drinking water fluoride, fluoride exposure dose, or fluorosis index in subjects across different IQ grades No statistical adjustment for covariates

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
Canada					
Green et al. (2019)[g]	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 urinary fluoride 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) or boys and girls combined (adjusted β = −1.95; 95% CI: −5.19, 1.28 and adjusted β = −1.24; 95% CI: −4.88, 2.40, 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 associations observed between measures of fluoride and verbal IQ Adjusted for sex, city, HOME score, maternal education, race, and prenatal secondhand smoke exposure

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
Till et al. (2020)[g]	Cohort (prospective) 10 cities/ MIREC [398] Non-fluoridated [247] Fluoridated [151] Breastfed as infants [200] Formula-fed as infants [198]	Drinking water Mean (SD) For breastfed infants: 0.13 (0.06) mg/L in non-fluoridated areas and 0.58 (0.08) mg/L in fluoridated areas For formula-fed infants: 0.13 (0.05) mg/day in non-fluoridated areas and 0.59 (0.07) mg/L in fluoridated areas Infant fluoride intake Mean (SD) For breastfed infants: 0.02 (0.02) mg/day in non-fluoridated areas and 0.12 (0.07) mg/day in fluoridated areas For formula-fed infants: 0.08 (0.04) mg/day in non-fluoridated areas and 0.34 (0.12) mg/day in fluoridated areas Maternal urine during pregnancy	Children (ages 3–4 years)	IQ: full-scale, performance, and verbal using Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III)	Drinking water Breastfed infants: Lower (not significant) full-scale IQ (adjusted β = −1.34, 95% CI: −5.04, 2.38) per 0.5-mg/L increase in water fluoride concentration; significantly lower performance IQ (adjusted β = −6.19, 95% CI: −10.45, −1.94) Formula-fed infants: Significantly lower full-scale IQ (adjusted β = −4.40, 95% CI: −8.34, −0.46) per 0.5-mg/L increase in water fluoride concentration; significantly lower performance IQ (adjusted β = −9.26, 95% CI: −13.77, −4.76) Infant fluoride intake Breastfed: No results reported Formula-fed: Lower (not significant) full-scale IQ (adjusted β = −2.69, 95% CI: −709, 3.21) per 0.5-mg/L increase in fluoride intake from formula; significantly lower performance IQ (adjusted β = −8.76, 95% CI: −14.18, −3.34) Maternal urine during pregnancy+

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
		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			Lower (not significant) full-scale IQ (adjusted β = −1.08, 95% CI: −1.54, 0.47) per 0.5-mg/L increase in maternal urinary fluoride++; lower (not significant) performance IQ (adjusted β = −1.31, 95% CI: −3.63, 1.03)++ Lower (not significant) performance IQ (adjusted β = −1.50, 95% CI: −3.41, 0.43) per 0.5-mg/L increase in maternal urinary fluoride+++; significantly lower full-scale IQ (adjusted β = −2.38, 95% CI: −4.62, −0.27)+++ No association between verbal IQ scores and any measure of fluoride exposure +Maternal urinary fluoride analyzed as covariate in the drinking water and infant fluoride intake from formula models and not in an individual model ++After additional adjustment for drinking water and breastfeeding status +++After additional adjustment for infant fluoride intake from formula All models adjusted for maternal education, maternal race, age at IQ testing, sex, HOME total score, and secondhand smoke status in the child's home (separate analysis also adjusted for mother's urinary fluoride)

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
India					
Sudhir et al. (2009)	Cross-sectional Nalgonda District (Andhra Pradesh)/school children [1,000]	Drinking water Level 1: <0.7 mg/L Level 2: 0.7–1.2 mg/L Level 3: 1.3–4.0 mg/L Level 4: >4.0 mg/L	Children (ages 13–15 years)	IQ: Raven's Standard Progressive Matrices	Significant increase in mean and distributions of IQ grades (i.e., increase in proportion of children with intellectual impairment) with increasing drinking water fluoride levels No statistical adjustment for covariates
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 correlations between IQ grade and water (r = 0.534) and urinary (r = 0.542) fluoride levels; in adjusted analyses, significant increase in mean IQ grade (i.e., increase in proportion of children with intellectual impairment) with increasing urinary fluoride; no significant differences in the levels of urinary lead or arsenic in children with the different water fluoride exposure levels Covariates included in the analysis were not reported
Trivedi et al. (2012)	Cross-sectional Kachchh, Gujarat/school children (6th and 7th grades) [84]	Mean (SE) Low-fluoride villages: drinking water: 0.84 (0.38) mg/L Children's urine: 0.42 (0.23) mg/L High fluoride villages: drinking water: 2.3 (0.87) mg/L Children's urine: 2.69 (0.92) mg/L	Children (ages 12–13 years)	IQ: questionnaire prepared by Professor JH Shah (97% reliability rating)	Significantly lower mean IQ score in high fluoride villages (92.53 ± 3.13) compared to the low-fluoride villages (97.17 ± 2.54); differences significant for boys and girls combined, as well as separately No statistical adjustment for covariates

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Summary of IQ Results[b,c]
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 association between water fluoride and IQ score (adjusted β = −3.865 per 1-mg/L increase in water fluoride); CIs not reported); significantly higher mean 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 age, sex, child's education level, mother's education level, father's education level, and fluorosis intensity

ANOVA = analysis of variance; GM = geometric mean; HOME = Home Observation Measurement of the Environment; IQ = intelligence quotient; Q1, Q3 = first and third quartiles; SD = standard deviations; WASI = Wechsler Abbreviated Scale of Intelligence (Spanish version); WISC-IV = Wechsler Intelligence Scale for Children-Revised (as reported by Choi et al. 2015).

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

[b]Associations between IQ and fluoride levels were reported quantitatively, when possible. For studies with multiple analyses and results, the table summarizes key findings and is not a comprehensive summary of all findings. Results also indicate when a study reported no association between IQ and fluoride, provided as a qualitative statement of no association.

[c]See Figure A-1 through Figure A-8 for additional study results.

[d]Xiang et al. (2003a), Xiang et al. (2011), and Wang et al. (2012) are based on the same study population.

[e]Yu et al. (2018) and Wang et al. (2020b) are based on the same study population.

[f]Three additional publications based on a subsample (i.e., 50–60 children) of the larger Yu et al. (2018) cohort were identified (Zhao et al. 2019; Zhao et al. 2020; Zhou 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.

[g]Green et al. (2019) and Till et al. (2020) are based on the same study population.

Summary of Results

Overall Findings

The results from 18 of the 19 high-quality (low risk-of-bias) studies (3 ~~longitudinal~~ prospective cohort studies from 2 different study populations and 15 cross-sectional studies from 13 different study populations) that evaluated IQ in children provide consistent evidence that higher fluoride exposure is associated with lower IQ scores (see "Summary of IQ Results" in Table 6) (Bashash et al. 2017; Choi et al. 2015; Cui et al. 2018; Ding et al. 2011; Green et al. 2019; Rocha-Amador et al. 2007; Saxena et al. 2012; Seraj et al. 2012; Sudhir et al. 2009; Till et al. 2020; Trivedi et al. 2012; Wang et al. 2012; Wang et al. 2020b; Xiang et al. 2003a; Xiang et al. 2011; Yu et al. 2018; Zhang et al. 2015b). Only one study (Soto-Barreras et al. 2019) did not observe an association between fluoride exposure and IQ; however, results were not provided in a manner that allowed for a direct comparison with other low risk-of-bias studies (see Appendix E for details). A strength of the findings across 18 of 19 low risk-of-bias studies was the consistent association between higher fluoride exposure [e.g., represented by populations whose total fluoride exposure approximated or exceeded the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] and lower IQ scores among studies of varying study designs, exposure measures, and study populations. In studies that analyzed the sexes separately (n = 5 studies with 2 studies reporting on the same study population), consistent findings of lower IQ associated with fluoride exposure were generally reported for both sexes. There is some indication of differential susceptibility between sexes, but ultimately, due to too few high-quality studies that analyzed exposure and outcome by sex separately and a lack of consistent findings that one sex is more susceptible, it is unclear whether one sex is more susceptible to the effects of fluoride exposure than the other. The body of evidence from the 19 low risk-of-bias studies is described in further detail below. Prospective cohort studies are discussed first, as this study design can establish a temporal relationship between exposure and outcome, which would contribute to demonstrating causality and, therefore, providing the strongest evidence for an association between fluoride exposure during development and IQ in children.

Results by Study Design – Prospective Cohort Studies

As noted above, three ~~longitudinal~~ prospective cohort studies, conducted in Mexico and Canada, were identified and considered to reflect a low risk for bias. All three prospective cohort studies found an association between increasing maternal or child fluoride exposure and lower IQ in children (Bashash et al. 2017; Green et al. 2019; Till et al. 2020). Two of the studies (Green et al. 2019; Till et al. 2020) were based on the same Canadian study population, but one evaluated prenatal fluoride exposure and the other evaluated postnatal fluoride exposure. Green et al. (2019) included maternal urinary fluoride, maternal fluoride intake, and water fluoride concentrations, while Till et al. (2020) used fluoride intake from formula or water concentrations in formula-fed versus breastfed infants. Multiple analyses were conducted in each prospective study, and results by analysis for the three prospective studies are discussed below. In summary, although not every analysis found a statistically significant association, together the three studies provided consistent evidence that increasing maternal fluoride levels were associated with lower IQ scores in the children.

In the Early Life Exposures in Mexico to Environmental Toxicants cohort, Bashash et al. (2017) observed a statistically significant association (p-value = 0.01) between lower IQ scores in children and prenatal fluoride exposure measured by maternal urinary fluoride (measured during

Commented [A53]: The following two sentences reflect revisions to clarify that while Green *et al. (2019)* and Till *et al.* (2020) use the same study population, the exposure measures used are different between the two publications, thus warranting consideration as separate studies, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 13): A small point, but [redacted] think the description of 19 studies somewhat exaggerates the size of the body of evidence, since these studies were conducted in 15 study populations. For example, on page 36, it is unclear why the two articles by Green and Till should get double the weight (2 vs. 1 study) simply because the authors chose to publish 2 (vs. 1) articles.

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 decrease in IQ score [95% CI: −4.12, −0.59] in boys and girls combined (see Figure A-8). This study also reported an inverse association between IQ level and children's urinary fluoride levels (single spot urine sample); however, this specific result did not achieve statistical significance (a 0.5-mg/L increase of child urinary fluoride was associated with a 0.89-point decrease in IQ score [95% CI: −2.63, 0.85]) (Bashash et al. 2017).

In the Maternal-Infant Research on Environmental Chemicals cohort, consisting of 10 cities in Canada, Green et al. (2019) also reported inverse associations between IQ scores in children and multiple measures of prenatal fluoride exposure, including maternal urinary fluoride, maternal fluoride intake, and water fluoride concentrations. Green et al. (2019) observed a statistically significantly lower IQ for boys associated with maternal urinary fluoride averaged across trimesters (4.49-point decrease in IQ score [95% CI: −8.38, −0.60; p-value = 0.02] per 1-mg/L increase in maternal urinary fluoride); however, results were not significant in boys and girls combined (1.95-point decrease in IQ [95% CI: −5.19, 1.28]) and were positive but not significant in girls (2.40-point increase in IQ [95% CI: −2.53, 7.33]). Other measures of prenatal exposure (maternal fluoride intake or water fluoride concentrations) were associated with lower IQ scores in boys and girls combined; the authors found no significant effect measure modification between child sex and fluoride exposure in these analyses so they did not report boys and girls separately (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 decrease in IQ score of 3.66 points in boys and girls combined (95% CI: −7.16, −0.15; p-value = 0.04). Similarly, based on drinking water ~~fluoride~~ concentrations ~~for pregnant women from fluoridated areas (mean water fluoride levels of 0.59 ± 0.08 mg/L) versus pregnant women from~~ and ~~non-fluoridated areas (mean water fluoride levels of 0.13 ± 0.06 mg/L),~~ a 1-mg/L increase of fluoride in drinking water was ~~were~~ associated with a significant 5.29-point decrease in IQ score ~~per 1-mg/L increase in fluoride~~ in both boys and girls combined (95% CI: −10.39, −0.19; p-value <0.05) (Green et al. 2019).

Commented [A54]: This sentence reflects revisions to clarify if results were statistically significant, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, Page 13): Please note when the result is not statistically significant and likely due to small sample sizes (e.g. discussion of the Green et al paper on page 37). Also for that paper, the results seem to be different by biological sex, an example of effect modification that would be expected for a neurotoxin.

Commented [A55]: Change made in response to the [redacted] Reviewer comment below; see DocB2_Monograph for detailed response:

Reviewer comment (DocB2_Monograph, Page 5): Is this actually observed (a 1 mg/L difference in fluoride concentrations leading to a 5.29 point decrease in IQ), or is this a predicted hypothetical effect from a model? If this is a modeled result rather than an observed result, should this be stated more clearly? Same comment may apply broadly.

In a study of the same study population as Green et al. (2019) that used fluoride intake from formula or water concentrations in formula-fed versus breastfed infants, Till et al. (2020) observed significantly lower performance IQ scores with higher fluoride regardless of the comparison used (p-values ≤0.004). They did not observe any association with verbal IQ, and full-scale IQ was only significantly lower in formula-fed infants using water fluoride concentrations as the exposure measure (p-value = 0.03). Breastfed infants and fluoride intake from formula also showed inverse associations but were not significant.

Taken together, the three prospective cohort studies (based on two North American study populations) indicate consistency in results across different types of analysis and across two study populations that higher fluoride exposure during development is associated with lower IQ scores.

Results by Study Design – Cross-sectional Studies

As with the prospective cohort studies, the cross-sectional studies reported a consistent association between fluoride exposure and lower IQ scores in children. Fifteen of the 16 low risk-of-bias cross-sectional studies [i.e., all with the exception of Soto-Barreras et al. (2019)]

consistently demonstrate that exposure to fluoride is associated with lower IQ scores. Fourteen of these 15 studies [with the exception of Cui et al. (2020)] reported significant associations.

Cross-sectional studies can have limitations, as the study design often cannot ensure that exposure preceded outcome. This uncertainty reduces confidence in study findings compared with prospective cohort studies—which, by design, establish that exposure occurred prior to outcome—and is captured in the outcome assessment. In some cases, cross-sectional studies do provide indicators of prior exposure (e.g., prevalence of dental fluorosis, limiting study populations to subjects who lived in the same area for long periods of time). Evidence that exposure occurred prior to the outcome of interest increases the confidence in results and any potential association reported in these studies. Of the 16 low risk-of-bias cross-sectional studies, 12 established that exposure preceded the outcome assessment (Choi et al. 2015; Ding et al. 2011; Rocha-Amador et al. 2007; Saxena et al. 2012; Seraj et al. 2012; Soto-Barreras et al. 2019; Sudhir et al. 2009; Wang et al. 2012; Wang et al. 2020b; Xiang et al. 2003a; Xiang et al. 2011; Yu et al. 2018). Five studies from different study populations indicated that a large portion of the exposed children had dental fluorosis (ranging from 43% to 100%) at the time of assessment (Choi et al. 2015; Ding et al. 2011; Seraj et al. 2012; Sudhir et al. 2009; Yu et al. 2018). 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. Nine studies from six study populations (including Yu et al. (2018) and Sudhir et al. (2009) listed above) excluded subjects who had not lived in the study area for a specified period of time, sometimes since birth (Rocha-Amador et al. 2007; Saxena et al. 2012; Soto-Barreras et al. 2019; Sudhir et al. 2009; Wang et al. 2012; Wang et al. 2020b; Xiang et al. 2003a; Xiang et al. 2011; Yu et al. 2018). Because these areas were generally known to be fluoride-endemic for long periods of time, it can generally be assumed that in these nine studies, exposure occurred prior to the outcome. Taken together, 12 cross-sectional studies from 9 study populations provide indicators of prior exposure.

Results by Study Design – Cross-sectional Study Variations

Overall, the cross-sectional studies consistently provide evidence that higher fluoride exposure is associated with lower IQ scores in children. Several cross-sectional studies conducted multiple analyses (e.g., reported results for multiple exposure metrics, endpoints, subpopulations). Although some of these variations are heterogeneous and are not comparable across studies, the consistency of the results across multiple metrics contributes to the confidence in the data. Table 6 summarizes key results for each of the low risk-of-bias cross-sectional studies, and a few examples of the within-study variations in results are provided below.

Nine cross-sectional studies (from six study populations) assessed the association between IQ and multiple exposure measures (Choi et al. 2015; Rocha-Amador et al. 2007; Saxena et al. 2012; Wang et al. 2012; Wang et al. 2020b; Xiang et al. 2003a; Xiang et al. 2011; Yu et al. 2018; Zhang et al. 2015b). Lower IQ was consistently observed across exposure measures in these studies; however, Choi et al. (2015), a small pilot study (n = 51), did not achieve statistical significance in all results by exposure measure. Specifically, the authors reported a consistent association between all fluoride exposure measures assessed (drinking water, children's urine, and severity of fluorosis) and digit span measures (subtest of the WISC-IV omnibus IQ test); however, results were only statistically significant when fluoride exposure was based on moderate or severe dental fluorosis in children (see Figure A-7). Choi et al. (2015) also observed

Commented [A56]: This sentence, and several sentences throughout the monograph, were revised to further distinguish between the comparison group and the group(s) exposed to higher levels of fluoride, in response to the [redacted] Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 12): One key feature for confidence rating is 'comparison group used.' This needs to be discussed further since fluoride exposure may be pervasive in water supplies. If so, in studies including a comparison group, include the comparison and how it was determined. Cross-sectional studies using biomarkers as continuous variables can be very strong.

Commented [A57]: Change made in response to the [redacted] comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

Commented [A58]: This sentence reflects revisions to clarify that the consistency of the results across multiple metrics contributes to the confidence in the body of evidence, rather than "increase", in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 14): The Results about the cross-sectional studies (page 38) state that "the consistent results across multiple metrics increase our confidence in the data." Based on the appropriate description in the Methods on Page 21, upgrading based on Consistency "does not apply in this evaluation".

some variation in results by outcome assessed (i.e., square root transformed block design and digit span [forward, backward, and total]). It was the only cross-sectional study that did not provide a full IQ score but instead provided results by specific subtests. The study authors consistently observed an inverse association between fluoride exposure and results from the digit span subtest (which specifically assesses executive function); however, results from the block design (square root transformed), a subtest of the WISC-IV omnibus IQ test that specifically assesses visuospatial function, was not associated with fluoride exposure. Note that Rocha-Amador et al. (2009) also assessed visuospatial function, and the authors reported a significant association (p-value <0.001) between fluoride exposure and decreased visuospatial constructional ability using the Rey-Osterrieth Complex Figure (ROCF) Test. Ultimately, too few studies were identified that reported results by subtest of omnibus IQ tests or assessed domains other than IQ (e.g., visuospatial function) to examine or explain the variation by outcome observed in Choi et al. (2015). The only other studies that provided a breakdown of the full IQ score were the prospective cohort studies by Green et al. (2019) and Till et al. (2020), which provided results for full-scale IQ as well as results for performance and verbal IQ. In both of these studies, lower verbal IQ was not associated with fluoride exposure, but lower performance and full-scale IQ were associated with fluoride exposure. There are too few studies to evaluate whether there is a specific aspect of IQ testing that is affected by exposure to fluoride, but the studies nonetheless consistently provide evidence that fluoride exposure is associated with lower IQ.

Yu et al. (2018) reported an overall association between lower IQ and higher fluoride exposure across multiple analyses but observed some variation in IQ results by urinary exposure level. The authors reported inverse associations between IQ and children's medium- and high-range urinary fluoride levels (1.60–2.50 mg/L and 2.50–5.54 mg/L, respectively), although change in IQ score was greater in the medium-range group (2.67 points decrease [95% CI: −4.67, −0.68]) for every 0.5-mg/L increase of urinary fluoride than in the high-range group (0.84 points decrease [95% CI: −2.18, 0.50]) (see Figure A-7). No association was reported at low-range urinary fluoride levels (0.01–1.60 mg/L). Note that Yu et al. (2018) also reported an inverse association between IQ and drinking water fluoride levels at 3.40–3.90 mg/L (4.29-point decrease in IQ score [95% CI: −8.09, −0.48]) for every 0.5-mg/L increase in water fluoride; a 0.04-point decrease in IQ score [95% CI: −0.33, 0.24] was observed for 0.5-mg/L increase in water fluoride at levels of 0.20–3.40 mg/L). The variation by exposure level in urine could not be verified in the analysis of drinking water exposures because there were only two water exposure groups (low and high). In a second study (Wang et al. 2020b), authors conducted a categorical analysis using urinary fluoride quartiles with reported betas per quartile. As observed in Yu et al. (2018), there were decreasing trends in IQ within each quartile; however, unlike Yu et al. (2018), Wang et al. (2020b) observed a larger decrease in IQ with each increasing urinary quartile and observed similar results using water fluoride quartiles (Wang et al. 2020b). Note that Wang et al. (2020b) cannot be compared directly to Yu et al. (2018) for evaluation at the higher exposure levels because the two studies do not use the same categorical exposure ranges. Although additional studies may have looked at different exposure levels, none of these studies provided results in the same manner as Yu et al. (2018) and Wang et al. (2020b) (i.e., betas by exposure category). Instead, these other studies provided an overall beta or mean IQ scores by exposure level. Despite the noted variations among these studies, the overall results still consistently support an association between fluoride exposure and lower IQ.

Two studies (Cui et al. 2018; Zhang et al. 2015b) observed associations between lower IQ in children and exposure to fluoride, with variations in results in subpopulations of children with different polymorphisms (see Figure A-7). These were the only two studies that considered polymorphism as a sub-analysis. Cui et al. (2018) observed a significant association between log-transformed children's single spot urinary fluoride and lower IQ scores (2.47-point decrease in IQ scores [95% CI: −4.93, −0.01; p-value = 0.049] per ln-mg/L increase in urinary fluoride), and the association was strongest in subjects with a TT polymorphism (compared with children with a CC or CT polymorphism) in the dopamine receptor D2 (DRD2) gene (12.31-point decrease in IQ score [95% CI: −18.69, −5.94; p-value <0.001] per ln-mg/L increase in urinary fluoride), which, according to the authors, probably resulted in a reduced D2 receptor density (Cui et al. 2018). Similarly, Zhang et al. (2015b) observed a significant association between lower IQ scores and children's single spot urinary fluoride (2.42-point decrease in IQ scores [95% CI: −4.59, −0.24; p-value = 0.030] per 1-mg/L increase in urinary fluoride), and the association was strongest in subjects with a val/val polymorphism (compared with children who carried the heterozygous or homozygous variant genotypes [met/val or met/met]) in the catechol-O-methyltransferase (COMT) gene (9.67-point decrease in IQ score [95% CI: −16.80, −2.55; p-value = 0.003] per 1-mg/L increase in urinary fluoride).

Overall, the cross-sectional studies consistently support a pattern of findings that higher fluoride exposure is associated with lower IQ scores in children. Slight within-study variations occur that may be associated with study variables such as IQ domains or subsets of IQ tests in a few studies that conducted multiple analyses, but these variations are heterogenous and cannot be further explored with the available studies. Despite these few variations, the overall evidence of an association with lower IQ is apparent.

Commented [A59]: This sentence, and several sentences throughout the monograph, reflect revisions to use the terms 'effect,' 'association,' and 'correlation' most appropriately, in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 17): Page 39, last full paragraph includes the sentence, "Despite these few variations, the overall evidence of an effect on IQ is apparent." This reviewer suggests editing the word "effect" to "association" or "correlation," given that the included studies are all observational.

Exposure Measure and Study Population Factors

Low risk-of-bias studies provide consistent evidence that higher fluoride exposure is associated with lower IQ scores across studies using different exposure measures. In addition to water fluoride levels, studies measured fluoride exposure using single serum samples in children (Xiang et al. 2011; Zhang et al. 2015b), single spot urine samples in children (Cui et al. 2018; Ding et al. 2011; Rocha-Amador et al. 2007; Saxena et al. 2012; Wang et al. 2020b; Xiang et al. 2003a; Yu et al. 2018; Zhang et al. 2015b), and prenatal maternal urinary measures (Bashash et al. 2017; Green et al. 2019), all of which were demonstrated to be consistently associated with lower IQ scores (see Figure A-6, Figure A-7, and Figure A-8). Urine levels encompass all sources of fluoride exposure and provide a better measure of the totality of exposure. As noted previously, even though some studies measured single spot samples, which may not be representative of peak exposure, these studies generally provided evidence that fluoride exposure had been occurring for some time. The consistency in the results across studies that used different measures of fluoride exposure and different life stages at which fluoride was measured strengthens the body of evidence.

The low risk-of-bias studies consistently provide evidence that higher fluoride exposure is associated with lower IQ scores across studies of different study populations. These 19 high-quality studies represent diverse populations (n = 15 study populations) across 5 countries. Eighteen of the 19 studies conducted in Canada (n = 2), China (n = 10), India (n = 3), Iran (n = 1), and Mexico (n = 2) provide evidence that exposure to fluoride is associated with lower IQ scores; 1 study conducted in Mexico did not observe an association but reported results in a

manner that did not allow for a direct comparison with the other studies (see Appendix E for details). The overall consistency in the study results across study populations adds strength to the body of evidence.

Exposure Levels

As described in this section, the body of evidence for studies assessing the association between fluoride exposure and IQ in children consistently provides evidence of an association between higher fluoride exposure [e.g., represented by populations whose total fluoride exposure approximates or exceeds the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] and lower IQ in children; however, there is less certainty in the evidence of an association in populations with lower fluoride exposures. In the September 6, 2019, draft of this monograph, NTP conducted a qualitative analysis of children's IQ studies that 1) evaluated lower fluoride exposures (<1.5 mg/L) in drinking water and/or urine and 2) provided information to evaluate dose response (i.e., provided three or more fluoride exposure groups or a dose-response curve in their publication) in the lower fluoride exposure range. Nine low risk-of-bias studies met these criteria, which includes the three prospective cohort studies discussed in this section. Based on the qualitative review of these studies, the evidence of an association between fluoride exposure below 1.5 mg/L and lower IQ in children appeared less consistent than results of studies at higher exposure levels.

A draft quantitative dose-response meta-analysis was prepared and included in the September 16, 2020, draft monograph (NTP 2020). This meta-analysis is undergoing further refinement in preparation for separate publication and may further inform a discussion on the association between fluoride exposure levels and IQ in children.

Sex Considerations

Recent literature suggests that adverse neurodevelopmental effects of early-life exposure to fluoride may differ depending on timing of exposure and sex of the exposed subject. In a review of the human and animal literature, Green et al. (2020) concluded that, compared with females, male offspring appear to be more sensitive to prenatal but not postnatal exposure to fluoride, with several potential sex-specific mechanisms.

Sex differences were examined in five of the low risk-of-bias studies (in four study populations) (Green et al. 2019; Trivedi et al. 2012; Wang et al. 2012; Wang et al. 2020b; Xiang et al. 2003a). In general, sex differences were difficult to assess for trends within different study populations because few studies in the body of evidence analyzed exposure and stratified results by sex. Although these five studies reported IQ scores separately for boys and girls, only two of these studies analyzed fluoride exposure for boys and girls separately (Green et al. 2019; Wang et al. 2020b), which is essential for evaluating whether a differential change in IQ by sex may be related to higher susceptibility in one sex or higher exposure in that sex. The remaining three studies stratified results by sex (Trivedi et al. 2012; Wang et al. 2012; Xiang et al. 2003a), but the analyses were based on area-level exposure data (e.g., low-fluoride village compared with high fluoride village) and not drinking water or urinary fluoride concentrations. In the five studies that reported results by sex separately, consistent findings of lower IQ associated with higher fluoride exposure were generally reported for both sexes. There was some variation in the results between sexes across study populations and exposure measures, but there is insufficient

Commented [A60]: This paragraph was added to clarify that recent publications had reported sex differences and describe potential sex-specific mechanisms in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 18): Page 40, "Gender considerations": Is there some biological plausibility that there would be sex differences in the relationship between fluoride exposure and neurocognitive outcomes. The term "susceptibility" is used several times, but it is unclear what that means. It seems to imply a biological reason, but it is unclear whether mechanistic evidence is supportive of that (or if gender differences actually represent some sort of residual confounding).

Commented [A61]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

evidence to determine whether one sex is more susceptible to the effects of fluoride exposure than the other.

Green et al. (2019) observed a significant inverse association between maternal urinary fluoride levels and IQ scores in boys (p-values ≤0.04) but not girls in a Canadian population. Green et al. (2019) did not find any sex differences in the association between IQ and water fluoride concentrations. Wang et al. (2020b) evaluated Chinese boys and girls separately and combined and observed statistically significant decreasing trends in IQ in all groups by urinary fluoride quartiles (p-values for trend ≤0.035) (see Figure A-7). Similarly, when evaluated as a continuous variable, spot urinary fluoride levels (per 1-mg/L increase) were significantly associated with lower IQ scores in girls (−1.379 [95% CI: −2.628, −0.129; p-value = 0.031]), boys (−1.037 [95% CI: −2.040, −0.035; p-value = 0.043]), and in the sexes combined (−1.214 [95% CI: −1.987, −0.442; p-value = 0.002]). According to water fluoride quartiles, Wang et al. (2020b) found that there was a significant trend in the sexes combined, although the decreasing trend in boys and girls separately did not achieve statistical significance (p-values = 0.077 and 0.055, respectively). When water fluoride levels were evaluated as a continuous variable (per 1-mg/L increase), there were significant associations with lower IQ scores in girls (−1.649 [95% CI: −3.201, −0.097]; p-value = 0.037), boys (−1.422 [95% CI: −2.792, −0.053; p-value = 0.042]), and the sexes combined (−1.587 [95% CI: −2.607, −0.568]; p-value = 0.002).

The remaining three studies that reported results by sex-based comparisons of areas of high and low urinary or water fluoride did not report exposure levels separately for boys and girls, which decreases the utility of the data to evaluate differential susceptibility by sex. Trivedi et al. (2012) observed significantly lower IQ in children in high fluoride Indian villages compared with low-fluoride villages with decreases observed in boys and girls separately or combined (p-values ≤0.05) (see Figure A-2). Xiang et al. (2003a) and Wang et al. (2012) provide data on the same study population in China. There was a significantly lower IQ in the high fluoride area compared with the low-fluoride area in boys and girls separately and in the sexes combined (p-values <0.01), although the difference was greater in girls. Because fluoride exposure was not analyzed for boys and girls separately, it is unclear whether the greater change in IQ scores in girls could be attributed to higher susceptibility to fluoride exposure or differences in fluoride exposure by sex.

In summary, it is unclear whether one sex is more susceptible to the effects of fluoride exposure than the other due to the limited number of studies that analyzed exposure and outcome by sex and the lack of a consistent pattern of findings that one sex is more susceptible. Green et al. (2019) did not observe an association between maternal urinary fluoride levels and IQ scores in girls but did observe a significant association in boys. Although this is an indication of higher sensitivity in boys in this analysis, the authors did not detect this sex difference using other measures of prenatal exposure (maternal fluoride intake or water fluoride concentrations). Wang et al. (2020b) and Trivedi et al. (2012) reported statistically significant associations in both boys and girls without indication that one sex may be more susceptible. Although Xiang et al. (2003a) and Wang et al. (2012) reported a greater change in IQ in girls than boys, the studies used area-level exposure data, and the authors did not determine whether fluoride exposure differed in boys versus girls. Therefore, it is unclear whether this differential result by sex is an indication of higher susceptibility in girls or whether it could be explained by a difference in exposure by sex. Overall, there are too few studies that analyzed exposure and outcome by sex separately to properly evaluate whether there is differential susceptibility to fluoride exposure by sex, and

results from the five low risk-of-bias studies that do evaluate sex differences indicate that there is no consistent difference by sex across the different study populations.

Summary of Key Findings for Low Risk-of-bias Children's IQ Studies

In summary, the high-quality studies (i.e., studies with low potential for bias) consistently demonstrate lower IQ scores with higher fluoride exposure [e.g., represented by populations whose total fluoride exposure approximates or exceeds the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)]. The consistency in association is observed among studies of varying study designs, exposure measures, and study populations. Although some studies that conducted multiple analyses observed within-study variations in results (e.g., differences between subsets of IQ tests), these variations were unique to individual studies and did not detract from the overall consistency in the findings that higher fluoride is associated with lower IQ scores.

Commented [A62]: This sentence reflects revisions to additional context for the use of the WHO Guidelines for Drinking -Water Quality value of 1.5 mg F/L, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 15):
Missing is a clear description or analysis across studies of what constitutes higher exposure levels that are associated with lower IQ. Page 40 starts a description of Exposure Levels, but lacks any quantitative description of high (or low) exposure. While [redacted] understand that a better analysis may arise from the future meta-analysis, there should be enough data in Table 6 to allow a more coherent summary of exposure level thresholds analyzed.
The *Summary of Key Findings for Low Risk-of-bias Children's IQ Studies* on Page 42 (and again on page 48) describes higher exposure as ⩾1.5 mg/L, but other than a mention of the 2016 report, this threshold is not described or presented in the Results. To the reader, this threshold is unsupported by the included studies.

High Risk-of-bias IQ Studies

The results from 53 studies with high potential for bias that evaluated IQ in children also consistently provide supporting evidence of decrements in IQ associated with exposures to fluoride. Forty-six of the 53 studies reported an association between high fluoride exposure and lower IQ scores in children.

Risk of Bias for IQ Studies in Children

The confidence in the human body of evidence was based on studies with the lowest potential for bias. A total of 19 studies on IQ in children had little or no risk-of-bias concerns, representing a relatively large body of evidence for low risk-of-bias studies (i.e., 15 study populations across 5 countries evaluating more than 7,000 children). These 19 studies are considered low risk of bias because they were rated probably low or definitely low risk of bias for at least two of the three key risk-of-bias questions and did not have any other risk-of-bias concerns that would indicate serious issues with the studies. Thirteen of the 19 studies were rated definitely low or probably low risk of bias for all risk-of-bias questions, and the remaining 6 studies were rated probably high risk of bias for a single question that was judged to have minimal impact on overall potential for bias. None of the 19 studies had a rating of definitely high risk of bias for any question. Risk-of-bias ratings for individual studies for all questions are available in Figure D-1 through Figure D-4, with risk-of-bias ratings for IQ studies in children available in Figure D-5 through Figure D-8 and Appendix E. Although the low risk-of-bias studies had minimal or no concerns, the studies with high overall potential for bias had 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. The key risk-of-bias questions are discussed below.

Confounding for IQ Studies in Children

Low Risk-of-bias Studies

As discussed above, there are 19 studies considered to have low risk of bias when assessed across all risk-of-bias domains. Sixteen of the 19 low risk-of-bias studies [i.e., all with the exception of Cui et al. (2020), Ding et al. (2011), and Soto-Barreras et al. (2019)] were considered to have low potential for bias due to confounding because the authors addressed the three key covariates for all studies (i.e., age, sex, and socioeconomic status) through study design

or analysis. Other important covariates, including health factors, smoking, and parental characteristics, were also addressed in many of the low risk-of-bias studies (see Figure 6).

Co-exposures to arsenic and lead were not considered a concern in 18 of 19 low risk-of-bias studies [i.e., all except for Soto-Barreras et al. (2019)] because the studies addressed the potential co-exposures, the co-exposures were not considered an issue in the study population, or the impact of the potential bias on the results was not a concern. Fifteen of 19 low risk-of-bias studies either addressed potential bias related to co-exposure to arsenic through study design or analysis or co-exposure to arsenic was unlikely in the study area. All 15 studies observed an association between lower IQ and higher fluoride exposure. Co-exposure to arsenic was not accounted for in the remaining four low risk-of-bias studies and was the main potential concern in these studies; however, three of these studies (Wang et al. 2012; Xiang et al. 2003a; Xiang et al. 2011) were still considered low risk of bias for confounding because although arsenic was observed in the water in the low-fluoride (and not the high-fluoride) comparison areas, which would bias the association toward the null, an association was still observed. In this case, the lack of adjustment for arsenic strengthens the evidence for an association and does not represent a potential concern. The other study did not address arsenic co-exposure and, as noted above, was conducted in an area that had potential for arsenic exposure to occur (Soto-Barreras et al. 2019); it is also the only low risk-of-bias study that did not observe an association between lower IQ and higher fluoride exposure (see Appendix E for further discussion of the risk-of-bias concern regarding arsenic for this study). Although Soto-Barreras et al. (2019) did not discuss arsenic, there is no direct evidence that arsenic was present in the study area. Fourteen studies accounted for co-exposure to lead through study design or analysis, and all observed an association between lower IQ and fluoride exposure. Five studies did not consider co-exposure to lead; however, for all of these studies, co-exposure to lead 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 (Cui et al. 2018; Cui et al. 2020; Soto-Barreras et al. 2019; Till et al. 2020; Trivedi et al. 2012).

There is considerable variation in the specific covariates considered across the 19 low risk-of-bias studies. The consistency of results across these studies suggests that confounding is not a concern in this body of evidence. Each of the 18 low risk-of-bias studies that observed an association between fluoride and IQ (see Summary of Results section above) considered a unique combination of covariates. The findings of these studies consistently provide evidence of an association between lower IQ in children and exposure to fluoride regardless of the inclusion or absence of consideration of any one or combination of covariates of interest. For example, maternal or family member smoking was addressed in 7 of the 19 low risk-of-bias studies, and this did not appear to affect the conclusions. All 7 studies that accounted for smoking found evidence of an association between fluoride exposure and lower IQ scores as did 11 of the 12 studies that did not account for smoking. Similarly, all 16 studies that addressed the three key covariates (age, sex, SES) (16 of 16 studies) and two of the three studies that did not fully account for them also found evidence of an association between fluoride exposure and lower IQ scores. In summary, when considering the impact of each covariate (or combinations of covariates) on the consistency of results, no trends are discernable that would suggest that bias due to confounding has impacted or would explain the consistency in findings across the body of evidence that fluoride exposure is associated with lower IQ in children.

Commented [A63]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

Commented [A64]: The following two sentences reflect revisions to further explain why the concern over co-exposure to arsenic in Soto-Barreras et al. (2019) would not result in the study being considered high risk of bias overall, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 17): It is unclear why Soto Barreras is considered to be low risk of bias (overall) if they did not account for arsenic in a high-exposure area. This seems like a major flaw. [redacted] did not find any description in the main part of the results (pages 28-41) that discuss this study and why it's included. Although, there's the unclear, unreferenced statement (page 36) that "Only one study did not observe evidence of an association between fluoride exposure and IQ; however, results were not provided in a manner that allowed for a direct comparison with other low risk-of-bias studies."

Commented [A65]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer Comment (DocF_Monograph, pages 5 - 11) : High?

Five of the low risk-of-bias studies confirmed the robustness of the results by conducting sensitivity analyses (Bashash et al. 2017; Green et al. 2019; Till et al. 2020; Wang et al. 2020b; Yu et al. 2018), and none of the sensitivity analyses adjusting for additional covariates 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. Green et al. (2019) reported 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 with the primary analyses. Wang et al. (2020b) found the results of the sensitivity analysis to be the same as the results from the primary analysis. Till et al. (2020) observed that adjusting for maternal urinary fluoride levels, as a way to consider postnatal exposure, had little impact on the results.

Among the 19 low risk-of-bias studies, three were identified that have potential for bias due to confounding (Cui et al. 2020; Ding et al. 2011; Soto-Barreras et al. 2019). This was mainly due to a lack of details on covariates considered key for all studies (i.e., age, sex, and SES). See Appendix E for further discussion of the risk-of-bias concerns regarding confounding for individual studies. Although these three studies have some potential for bias due to confounding, they are considered to be low risk of bias overall because they have low potential for bias for the other two key risk-of-bias questions (exposure characterization and outcome assessment), and no other major concerns for bias were identified. Consistent with the 16 studies that adequately addressed confounding, two of these three studies also provide evidence of an association between fluoride exposure and lower IQ scores in children.

Taken together and considering the consistency in the results despite the variability across studies in which covariates were accounted for, bias due to confounding is not considered to be a concern in the body of evidence. The potential for the consistency in results to be attributable to bias due to confounding in the 19 low risk-of-bias studies is considered low.

Study (Location)[a]	Potential Covariates Considered[b]																Notes	Reported Association with Fluoride[e]
	Subject Characteristics				Other Exposures					Socioeconomic Factors		Parental Characteristics				Other[c]		
	Age	Sex	Race/Ethnicity	Health Factors[c]	Arsenic	Smoking	Iodine	Lead	Other[c]	SES[d]	Caregiving Environment (e.g., HOME score)	Demographics[c]	Reproductive Factors[c]	Health Factors[c]	IQ			
Overall RoB Rating for Confounding: Probably Low																		
Bashash 2017 (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
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 2012 (India)	√	√	–	–	√	–	√	–	–	√	–	–	–	–	–	–		Yes
Wang 2012 (China)	√	√	–	√	–	–	√	√	–	√	–	–	–	√	–	√	Health: medical conditions Other: transportation, natural environment, and lifestyle	Yes
Wang 2020b (China)	√	√	–	√	√	√	√	√	–	√	–	–	√	–	–	√	Health: subject BMI, exclusions based on diseases affecting intelligence, history of trauma or neurological disorders, positive screening test history Reproductive: low birth weight Other: alcohol consumption	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
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 hormone levels	Yes[f]
Ding 2011 (China)	√	–	–	–	√	–	√	√	–	–	–	–	–	–	–	–		Yes
Soto-Barreras 2019 (Mexico)	√	√	–	–	–	–	–	–	–	√	–	–	–	–	–	–		No

Figure 6. Important Covariates Considered in Low Risk-of-bias IQ Studies Conducted in Children

[a]Includes all low risk-of-bias IQ 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).
[b]Covariates represented here are those considered important for this evaluation. Depending on the specific study population, individual covariates may be considered a potential confounder, effect measure modifier, and/or co-exposure. See study details provided in HAWC (NTP 2019) for information on additional covariates.
Factors outlined in blue are key covariates for all studies (subject age, subject sex, SES) and arsenic (which is of particular importance to some study populations).

[A] √ indicates that a covariate was considered. Examples of what it means for a covariate to be "considered": it was adjusted for in the final model, it was considered in the model but not included in the final model because it did not change the effect estimate, it was reported to have the same distribution in both the exposed and unexposed groups, it was reported to not be associated with the exposure or outcome in that specific study population. For arsenic, a √ might also be used when arsenic 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 Appendix E [or HAWC (NTP 2019) for details. A hyphen (-) indicates that the factor was not considered.
[c]See the "Notes" column for additional details.
[d]Covariates considered measures of SES include SES scaled scores, household/family income, child education, caretaker/parental education, and occupation/employment.
[e]Extent of reported associations varies by study. "Yes" indicates that study authors provided evidence of an association between lower IQ scores and fluoride exposure.
[f]Study reported lower IQ scores with increasing fluoride exposure, but the results did not achieve statistical significance.

High Risk-of-bias Studies

Most high risk-of-bias studies (n = 53) considered important covariates to some degree through study design or analysis; however, when considering the full scale of potential concerns of bias due to confounding, all but three of these studies were rated probably or definitely high risk of bias. The majority of high risk-of-bias studies accounted for one or two of the three covariates considered key for all studies (age, sex, SES) but did not address all three and did not address other covariates considered important for the specific study population and outcome. Potential confounding related to important co-exposures (e.g., arsenic) was often not addressed in high risk-of-bias studies. In studies in which 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.

Despite the lack of adequate consideration of key covariates in the vast majority of high risk-of-bias studies, the results across most of these studies (46 of 53) consistently provide evidence of an association between fluoride exposure and IQ, supporting the results observed in the low risk-of-bias studies. This finding suggests that confounding is likely less of a concern for the body of evidence as a whole than for any individual study. Although the high risk-of-bias studies may have more potential for bias due to confounding compared with the low risk-of-bias studies, the consistent IQ findings across high and low risk-of-bias studies indicate that the results cannot be explained solely by potential bias due to confounding.

Exposure Characterization in IQ Studies

Low Risk-of-bias Studies

In general, there were few, if any, risk-of-bias concerns regarding exposure characterization in the low risk-of-bias studies. These studies mainly had individual exposure data based on urine or water measures with appropriate analyses. Although there are concerns related to using urine samples (see the Risk-of-bias Considerations for Human Studies section for details), the evidence suggests that urinary fluoride is a reasonable measure of exposure (Villa et al. 2010; Watanabe et al. 1995). 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 maternal urinary fluoride for creatinine did not substantially alter the observed association (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) included only participants with valid fluoride

Commented [A66]: This footnote reflects revisions to clarify what we mean by "consider", in response to related comments from [redacted]Reviewers listed below; see DocG_Monograph and DocJ_Monograph, respectively, for detailed response:

Reviewer comment (DocG_Monograph, page 18): It's unclear what is meant (in Table 6 and scattered throughout the Results) that there was "No statistical adjustment for confounders" but then in Figure 6 (and also in the text) all studies "consider" the potential confounders age, sex, and SES.

Reviewer comment (DocJ_Monograph, page 14): In general, it is difficult to understand how cross-sectional studies that adjusted for few, or no, confounders, employed somewhat indirect measures of fluoride exposure (or did not fully capture all sources of exposures to fluoride), or had concerns related to selection bias, were designated as "low risk of bias." If, for example, some confounders were accounted for in the design or analysis, other than statistical adjustment, it may be worth noting that on Table 6 (otherwise, it appears that many papers accounted for no confounders).

For example, Xiang, 2003a did not statistically adjust for any confounders. They did report some findings in relation to some of the confounders, but not to the extent that [redacted]would perceive them to have been fully accounted for. ...

Commented [A67]: This footnote was added to list all measures related to SES that were considered in the low risk-of-bias IQ-in-children studies, in response to related comments from [redacted]Reviewers listed below; see DocH_Monograph, DocI_Monograph, and DocK_Monograph, respectively, for detailed responses:

Reviewer comment (DocH_Monograph, page 5): RE: confounding and covariates. Recent thinking regarding confounding requires the use of directed acyclic graphs to define variables which are theoretically confounders (based on previous literature). Thus, some clarification is needed on how the set of three important confounders were selected, i.e. sex, child age and a measure of socioeconomic status. Indeed, based on literature from other potential neurotoxins (e.g. lead, polychlorinated biphenyls, phthalates) it seems as though child sex would be an effect modifying variable, not a confounder (child sex would not be related, for example to exposure status under any definition of confounding). Variables such as arsenic or lead exposure would be co- ...

Commented [A68]: This footnote was added to provide further clarity that parental education is captured under SES in Figure 6, in response to the [redacted]Reviewer comment below; see DocK_Monograph for detailed response:

Reviewer comment (DocK_Monograph, page 4): Additionally, it might be helpful to identify a limited set of confounder as required for evaluation. For example, those included in Figure 6 do not include all described in Table 6, and in fact m[a]y not present an important one: parental educational attainment.

measurements at all trimesters in their analysis. Other studies also measured urinary fluoride multiple times throughout pregnancy (Bashash et al. 2017). Some studies demonstrated correlations between urinary fluoride and fluoride in drinking water, fluorosis, or estimated dose based on drinking water concentrations and consumption (Choi et al. 2015; Ding et al. 2011; Green et al. 2019; Saxena et al. 2012; Yu et al. 2018; Zhang et al. 2015b). 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 used to correct for urine dilution or whether adjustments were made for dilution. Bashash et al. (2017) excluded exposure outliers and found that doing so did not substantively change the results. Taken together, these studies suggest that urinary fluoride is a reasonable measure of exposure despite some potential issues.

All but one low risk-of-bias study was rated probably or definitely low risk of bias for exposure assessment. Seraj et al. (2012) had potential exposure misclassification and was rated probably high risk of bias for exposure assessment. Villages were categorized as normal (0.5–1 ppm), medium (3.1 ± 0.9 ppm), or high (5.2 ± 1.1 ppm) based on average fluoride content in drinking water in varying seasons over a 12-year period. Mild fluorosis observed in children in the normal fluoride level group indicates that there may have been higher exposure in this group at some point in the past; however, this would bias the results toward the null, and the children in the normal fluoride group had a significantly higher IQ score compared with the medium and high fluoride groups (p-value = 0.001). There were also significant associations between lower IQ scores and fluorosis intensity (p-value = 0.014) and water fluoride concentration when evaluated as a continuous variable (p-values <0.001). Although there is potential for exposure bias, the apparent exposure misclassification and inclusion of children with higher fluoride exposure in the normal group indicate that the association may be greater than what was observed in this study.

High Risk-of-bias Studies

A frequent, critical limitation among the high risk-of-bias studies was lack of information regarding exposure or poor exposure characterization. Many of the high risk-of-bias studies compared only subjects living in two regions with differing levels of fluoride exposure, and although 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 whether 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 with non-endemic areas, or high-fluoride areas were compared with low-fluoride areas, with no other information provided on fluoride levels in the areas (Li et al. 2003 [translated in Li et al. 2008c]; Ren et al. 1989 [translated in Ren et al. 2008]; Sun et al. 1991). Although living in an area endemic for fluorosis could be an indicator of exposure, these studies did not specify whether the study subjects themselves had fluorosis. Another study used only dental fluorosis as a measure of fluoride exposure in subjects who were all from an endemic area with similar drinking water fluoride levels (Li et al. 2010). In one case, multiple sources of fluoride exposure were assessed separately without properly controlling for the other sources of exposure, which could bias the results (Broadbent et al. 2015). Broadbent et al. (2015) assessed fluoride exposure in three ways: use of community water in a fluoridated area

versus a non-fluoridated area, use of fluoride toothpaste (never, sometimes, always), or use of fluoride tablets prior to age 5 (ever, never). The same children were used for each analysis without accounting for fluoride exposure through other sources. For example, there were 99 children included in the non-fluoridated area for the community water evaluation, but there is no indication that these 99 children were not some of the 139 children that had ever used supplemental fluoride tablets or the 634 children that had always used fluoride toothpaste. Therefore, comparing fluoridated areas to non-fluoridated areas without accounting for other sources of exposure that might occur in these non-fluoridated areas would bias the results toward the null.

Outcome Assessment for IQ Studies

Low Risk-of-bias Studies

The low risk-of-bias studies have few concerns regarding outcome assessment. All 19 low risk-of-bias studies used appropriate methods for measuring IQ in the study population being assessed, and blinding of outcome assessors was not a concern in 18 of the 19 studies [i.e., all low risk-of-bias studies except Sudhir et al. (2009)]. Fourteen of these 18 studies reported blinding of the outcome assessors, or correspondence with the study authors confirmed that it was not likely an issue. For the remaining 4 of the 18 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 in the general population studies. One IQ study (Sudhir et al. 2009) had concerns for potential bias in the outcome assessment due to lack of information to determine whether blinding at the time of the outcome assessment was a concern (see Appendix E for details).

High Risk-of-bias Studies

Among the studies with high risk of bias, the main limitation in the outcome assessment was the lack of reporting on blinding of the outcome assessor (i.e., 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.

High risk-of-bias studies were mainly carried out in two separate populations without information provided that the tests were conducted in a central location. In many cases, the methods indicated that the tests were conducted at the schools in the study area (indicating that there was likely knowledge of exposure). 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).

Confidence Assessment of Findings on IQ in Children

We conclude that there is moderate confidence in the body of evidence that higher fluoride exposure is associated with lower IQ in children. This confidence rating was reached by starting

Commented [A69]: This sentence, and many sentences throughout the monograph, reflect revisions to further clarify the studies to which we are referring, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 18): The Outcome Assessment for IQ Studies section (page 48) is unclear. This problem occurs in much of the write up, where it is unclear what studies are being referred to. It states that 18 of 19 studies were low risk ("used appropriate methods for measuring IQ"), but does not indicate which study did not use appropriate methods or what the problem is. At the end of the paragraph there's a sentence about Sudhir not reporting blinding, but the paragraph starts by saying that "blinding of outcome assessors was not a concern).

Commented [A70]: This paragraph reflects revisions to rearrange the text to start with the initial confidence rating and how initial confidence rating is determined, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 20): A clearer statement, up front, is needed that the starting point for confidence is "moderate" and why this is the case. [redacted] think you're trying to say this with "The initial moderate confidence rating in the body of evidence" on page 48, but this sentence is unclear. [redacted] still unsure if "initial" here means where the GRADE confidence rating starts before assessing the evidence. Why start at moderate? The Methods section does not describe this concept.

with an initial confidence rating based on key study design features of the body of evidence and then considering factors that may increase or decrease the confidence in that body of evidence. The initial moderate confidence rating is based on 15 of the 19 low risk-of-bias studies that have 3 of the 4 key study design features shown in Figure 1 (i.e., exposure occurred prior to outcome, individual-based outcomes were evaluated, and a comparison group was used). Three of these studies were prospective cohort studies, and 12 were cross-sectional studies that provided evidence of long-term, chronic fluoride exposure prior to outcome measurement.

There are nine factors to consider for increasing or decreasing the confidence in the body of evidence (provided in Figure 1). Discussion of each of these factors in the body of evidence on fluoride exposure and IQ in children is presented below.

- **Risk of bias**: Only studies that were considered to have low risk of bias were included in the moderate confidence rating; therefore, there was no downgrade for risk-of-bias concerns.
- **Unexplained inconsistencies:** The ~~data are~~direction of the association is consistent in the majority of studies, and there was no downgrade for this factor. Eighteen of the 19 low risk-of-bias studies reported associations between higher fluoride levels and lower IQ scores in children. These studies were conducted in 5 different countries on more than 7,000 children from 15 different study populations. There is consistency in ~~results~~ the direction of the association across prospective and cross-sectional study designs. There is also consistency in ~~results~~ the direction of the association across studies using different fluoride exposure measures, including urinary and drinking water fluoride. The one study that did not observe an association did not provide results in a comparable manner and therefore this body of evidence is not considered to have unexplained inconsistencies.
- **Indirectness:** IQ in humans is a direct measure of the association of interest; therefore, no adjustment in confidence is warranted.
- **Imprecision:** There is no evidence of serious imprecision that would warrant a downgrade. Eighteen low risk-of-bias studies reported lower IQ with higher fluoride, and no issues with imprecision were identified to challenge the significance of the ~~effect~~ response estimates.
- **Publication bias:** There is no strong evidence of publication bias; therefore, no downgrade was applied for publication bias. Two published meta-analyses (Choi et al. 2012; Duan et al. 2018) did not indicate strong evidence of publication bias. The draft meta-analysis conducted by NTP in the September 16, 2020, draft monograph found no publication bias among the low risk-of-bias studies (NTP 2020). Among high risk-of-bias studies, adjusting for publication bias using the trim-and-fill analysis estimated that, in the absence of publication bias, the inverse direction of association and statistical significance remained, thus indicating that there was no need to downgrade for publication bias.
- **Large magnitude of effect size:** Although some individual studies indicated a large magnitude of effect size, the magnitude of effect was not the same across all studies. Therefore, the overall data would not support an upgrade due to a large magnitude of

Commented [A71]: This bullet reflects revisions to further characterize the consistency in results in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 19): Page 48-49, Assessment of Unexplained Inconsistencies: While there is some consistency in findings suggesting that increased exposure to fluoride is associated with lower IQ, many studies reported mixed results (generally reporting a mix of inverse associations and null findings). How were these mixed findings taken into consideration when evaluating unexplained inconsistencies?

Commented [A72]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page15): Imprecision clearly evident upon visual inspection of Figures A1 through A5, which frequently shows wide and overlapping confidence intervals. Therefore, this discussion item should be revised.

Commented [A73]: This sentence was revised to clarify that serious imprecision as described in the protocol is required to downgrade, which was not found for this body of evidence, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, Page 12): Regarding the GRADE assessment of imprecision, please provide more detail about your thresholds between precise and imprecise. For example, what 95% CI would indicate imprecise?

- **Dose response:** Evidence of an exposure-response relationship that could justify an upgrade to the confidence in the body of evidence is not presented in this monograph. While the overall findings qualitatively appear less clear in the lower exposure range, many of the studies that provide data to evaluate exposure response were judged to be high risk of bias. The meta-analysis conducted in association with this systematic review further informs this issue and will be ~~published seperately~~refined in preparation for a separate publication.
- **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 were significantly lower IQ scores in the area with high fluoride. The remaining studies do not provide enough information to consider whether residual confounding occurred for the body of evidence. Note that parental IQ has the potential to be an important factor when considering residual confounding based on likely correlations between parental IQ and children's IQ; however, there is not sufficient evidence that parental IQ is associated with water fluoride content. Taken together, the overall data would not support an upgrade due to residual confounding.
- **Consistency:** The consideration of a potential upgrade for consistency in the methods is primarily for non-human animal evidence, where it would be applied to address increased confidence for consistent effects across multiple non-human animal species. For human evidence, it is generally not applied, and the data would only be considered in deciding whether to downgrade for unexplained inconsistency. Therefore, no upgrade is applied for consistency.

As described above, there are no changes in confidence rating based on any of the possible upgrade or downgrade factors. The magnitude of effect size and the overall strength and quality of the human literature base provide moderate confidence in the body of evidence that higher exposure to fluoride is associated with lower IQ in children (see the Discussion section for strengths and limitations of the evidence base). Note that additional, well-designed prospective cohort studies with individual-level exposure data and outcome measures could provide increased confidence in the association between fluoride exposure and lower IQ in children.

Other Neurodevelopmental or Cognitive Effects in Children

Low Risk-of-bias Studies

Overview of Studies

Nine low risk-of-bias studies (three prospective cohort and six cross-sectional studies) evaluated the association between fluoride exposure and cognitive neurodevelopmental effects other than IQ in children. These nine studies were conducted in multiple study populations in three countries, specifically:

- three were conducted in three areas of China on three study populations,
- four were conducted in two areas of Mexico on three study populations, and
- two were conducted in Canada using the same study population.

Commented [A74]: This bullet reflects revisions to limit the description of the dose-response relationship, since the dose-response meta-analysis was removed from the monograph, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 19): Similarly, the description around "Dose-response" on page 49 is not clearly supported by the text of the Results section. There is no clear dose-response section of the Results where related findings are described. The Results text mostly summarizes as "high" or "exposure" or in some instances association with a 1-mg/L increase or the equivalent.

Commented [A75]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, page 13): Would revise because it is being submitted for publication and also [redacted] would not link it to further informing this since there are still questions about the meta-analysis.

Commented [A76]: This bullet reflects revisions to remove text on the consistency across studies, as this was addressed in the preceding Unexplained Inconsistencies bullet, in response to the [redacted] Reviewer comment below; see
DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 19): The "Consistency" section on page 49 should not discuss the consistency across studies. This was addressed in Unexplained Inconsistencies on the prior page. Do not confuse the two issues for the reader.

Commented [A77]: This sentence reflects revisions to include information on counts of studies per study design, in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 12): In general, it would be helpful to the reader to describe the included study designs earlier in the respective results sections (i.e., along with the total #s of included articles, describing the # of cross-sectional, prospective cohort, etc is important).

There is considerable heterogeneity across studies, particularly in the different health outcomes evaluated and ages assessed. Most studies measured fluoride in the drinking water or urine (child or maternal) with one study using severity of dental fluorosis as an exposure measure in addition to drinking water and children's urine. Two of the studies were conducted on infants, with one evaluating effects within 72 hours of birth (Li et al. 2004 [translated in Li et al. 2008a]) and the other evaluating effects at 3 to 15 months of age (Valdez Jimenez et al. 2017). The remaining studies were conducted in children of varying ages, ranging from 4 to 17 years. Other cognitive neurodevelopmental outcomes assessed include neurobehavioral effects in infants, learning and memory impairment, and learning disabilities such as attention deficit hyperactivity disorder (ADHD). Few studies measured the same health outcomes, used the same outcome assessment methods, or evaluated the same age groups.

Table 7 provides a summary of study characteristics and key findings related to other cognitive neurodevelopmental outcomes and fluoride exposure for the nine low risk-of-bias studies. The different tests conducted and the populations on which the tests were conducted are also indicated in Table 7. Several of these studies conducted multiple analyses and reported results on multiple endpoints. The purpose of the table is to summarize key findings (independent of whether an association was found) from each study and is not meant to be a comprehensive summary of all results. For each study, results are summarized for each exposure measure assessed. Results from multiple analyses using the same exposure measure may not all be presented unless conflicting results were reported. See Appendix E for additional information on studies in Table 7, including strengths and limitations, clarifications for why they are considered to pose low risk of bias, and information regarding statistical analyses, covariates, exposure assessment, and outcome assessment.

Commented [A78]: This sentence was added to direct the reader to Appendix E for information on statistical methods and other suggestions for Table 6 in response to the [redacted]Reviewer comment below; see DocI_Monograph for detailed response:

Reviewer comment (DocI_Monograph, page 7): Table 6 could include the following: 1) statistical methods; 2) confounders, particularly exposure to other known neurotoxicants, and how they were measured; 3) might rename 'Assessment timing' to age of participants or just combine the information with the location/subject's column

Table 7. Studies on Other Neurodevelopmental and Cognitive Function in Children[a]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurobehavioral Outcome Summary[b]
China					
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) (subjects divided into high fluoride group and control group based on drinking water fluoride levels in place of residence); 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 covariates
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 any fluoride exposure measure Adjusted for age, sex, parity, illness before 3 years old, household income last year, and caretaker's age and education
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); no associations between urinary fluoride level and ADHD index or other behavioral measures Adjusted for age, sex, child's BMI, urinary creatinine, mother migrated, and father migrated

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurobehavioral Outcome Summary[b]
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]	Maternal urine Range: 0.16–8.2 mg/L (all trimesters) Drinking water Range: 0.5–12.5 mg/L (all trimesters)	Infants (ages 3–15 months)	Mental development index (MDI): Bayley Scales of Infant Development II (BSDI-II) Psychomotor developmental index (PDI): Bayley Scales of Infant Development II (BSDI-II)	Significant association between log10-mg/L maternal urinary fluoride and MDI score during first trimester (adjusted $\beta = -19.05$; SE = 8.9) and second trimester (adjusted $\beta = -19.34$; SE = 7.46); no significant associations between maternal urinary fluoride and PDI score; analyses of outcomes using drinking water fluoride not performed Adjusted for age, gestational age, marginality index, and type of drinking water
Bashash et al. (2017)[c]	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 association between maternal urinary fluoride and offspring GCI score (per 0.5-mg/L increase adjusted $\beta = -3.15$; 95% CI: −5.42, −0.87); associations with children's urine not significant Adjusted for gestational age; weight at birth; sex; 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

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurobehavioral Outcome Summary[b]
Bashash et al. (2018)[c]	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 (per 0.5-mg/L increase) 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; sex; 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
Canada					
Barberio et al. (2017b)[d]	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) per 1-µmol/L increase in unadjusted urinary fluoride when Cycle 2 and 3 were combined; no significant associations found between urinary fluoride and ADHD (only evaluated in Cycle 2); no significant associations found when using creatinine- or specific gravity-adjusted urinary fluoride Adjusted for age and sex, household income adequacy, and highest attained education in the household

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurobehavioral Outcome Summary[b]
Riddell et al. (2019)[d]	Cross-sectional General population/Canadian Health Measures Survey (Cycles 2 and 3) [3,745]	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)] Community water fluoridation status (yes or no) 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)]	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 age and sex, child's BMI, ethnicity, parental education, household income, blood lead, and smoking in the home

ADHD = attention-deficit/hyperactivity disorder; BMI = body mass index; GCI = General Cognitive Index; GM = geometric mean; HOME = Home Observation Measurement of the Environment; IQ = intelligence quotient; MSCA = McCarthy Scales of Children's Abilities; SD = standard deviation; WASI = Wechsler Abbreviated Scale of Intelligence (Spanish version); WISC-IV = Wechsler Intelligence Scale for Children-Revised (as reported by Choi et al. 2015); WRAML = Wide Range Assessment of Memory and Learning; WRAVMA = Wide Range Assessment of Visual Motor Ability.

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

[b]Associations between other cognitive neurodevelopmental outcomes in children and fluoride levels were reported quantitatively, when possible. For studies with multiple analyses and results, the table summarizes key findings and is not a comprehensive summary of all findings. Results also indicated when a study reported no association, provided as a qualitative statement of no association.

[c]Bashash et al. (2017) and Bashash et al. (2018) are based on the same study population.

[d]Barberio et al. (2017b) and Riddell et al. (2019) are based on the same study population.

Summary of Results

Overall Findings

Although discussed together in this section, various health outcomes were assessed in the nine low risk-of-bias studies of other neurodevelopmental outcomes, including neurobehavioral scores in infants (two studies), cognitive tests in children other than IQ (three studies), and ADHD or learning disabilities (four studies) in children. ~~Altogether, t~~The results from eight of nine low risk-of-bias studies (three prospective cohort studies and five cross-sectional studies from seven different study populations) provide evidence of significant associations between higher fluoride exposure and cognitive neurodevelopmental outcomes in children other than decrements in IQ (see Figure A-9 through Figure A-11) (Barberio et al. 2017b; Bashash et al. 2017; Bashash et al. 2018; Li et al. 2004 [translated in Li et al. 2008a]; Riddell et al. 2019; Rocha-Amador et al. 2009; Valdez Jimenez et al. 2017; Wang et al. 2020a). Only one cross-sectional study did not find a significant association between fluoride exposure and a measure of cognitive neurodevelopment (Choi et al. 2015).

Although there is heterogeneity in the outcomes assessed and a limited number of directly comparable studies, the data provide additional evidence (beyond the consistent evidence of an association between fluoride exposure and IQ) of an association between higher fluoride exposure and cognitive or neurodevelopmental effects. The body of evidence from the nine low risk-of-bias studies is described in further detail below and is grouped into outcome categories of studies that are most comparable.

Results in Infants

Two studies evaluated neurobehavioral effects in infants either shortly after birth or at 3 to 15 months of age (Li et al. 2004 [translated in Li et al. 2008a]; Valdez Jimenez et al. 2017). Both studies observed a significant association between higher fluoride exposure and lower neurobehavioral scores. In neonates (1–3 days old), the high fluoride group (3.58 ± 1.47 mg/L fluoride based on spot maternal urine collected just prior to birth) had significantly lower total neurobehavioral assessment scores (36.48 ± 1.09 versus 38.28 ± 1.10 in controls; p-value <0.05) and total behavioral capacity scores (10.05 ± 0.94 versus 11.34 ± 0.56 in controls; p-value <0.05) compared to the control group (1.74 ± 0.96 mg/L fluoride) as measured by a standard neonatal behavioral neurological assessment (NBNA) method (Li et al. 2004 [translated in Li et al. 2008a]). In infants 3 to 15 months of age, the Mental Development Index (MDI)—which measures functions including hand-eye coordination, manipulation, understanding of object relations, imitation, and early language development—was significantly inversely associated with maternal urinary fluoride in both the first and second trimesters (adjusted βs per log10-mg/L increase = −19.05 with standard error of 8.9 for first trimester [p-value = 0.04] and −19.34 with standard error of 7.46 for second trimester [p-value = 0.013]) (Valdez Jimenez et al. 2017). Note that this study did not find an association between maternal fluoride during any trimester and the Psychomotor Developmental Index (PDI), which measures gross motor development (adjusted βs = 6.28 and 5.33 for first and second trimesters, respectively; no standard errors provided) (Valdez Jimenez et al. 2017).

Results for Cognitive Tests Other Than IQ in Children

Three studies conducted tests on cognitive function in children that were not part of an IQ test (Bashash et al. 2017; Choi et al. 2015; Rocha-Amador et al. 2009). None of the studies

Commented [A79]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 17): In this section, this summary statement without further explanation is misleading. Elsewhere in this document the authors indicate that the data regarding ADHD effects contains significant heterogeneity regarding methods and outcomes and thereby precludes conclusions about ADHD and other attention-related disorders.

Commented [A80]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

Commented [A81]: This sentence and text throughout the monograph reflect changes to clarify that the associations are per log10-mg/L increase in fluoride exposure, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 18): Some of the associations are really quite large, e.g. adjusted betas of -19 in the study of Valedez Jimenez et al 2017, especially for the Bayley Scale. Such associations are either suspect or are not adjusted for the concentration of fluoride appropriately (maybe it is a log unit change). This needs to be clarified.

conducted the same tests, but two of the three studies (Bashash et al. 2017; Rocha-Amador et al. 2009) observed associations between fluoride exposure and lower test scores. The General Cognitive Index (GCI) of the McCarthy Scales of Children's Abilities (MSCA) in 4-year-old children was significantly inversely associated with maternal creatinine-adjusted urinary fluoride levels during pregnancy (collected during each trimester) (adjusted β per 0.5-mg/L increase = −3.15 [95% CI: −5.42, −0.87; p-value = 0.01] in a model adjusting for main covariates including gestational age, weight at birth, sex, maternal smoking, and indicators of socioeconomic status). The association remained even after adjusting for maternal bone lead (adjusted β per 0.5-mg/L increase = −5.63 [95% CI: −8.53, −2.72; p-value <0.01]) (Bashash et al. 2017) (see Figure A-11). Choi et al. (2015), however, evaluated cognitive function endpoints in addition to IQ and found no significant associations between concurrent log-transformed water or urinary fluoride levels and Wide Range Assessment of Visual Motor Ability (WRAVMA) scores, finger tapping test scores, and grooved pegboard test scores, although there were some significant associations based on degree of fluorosis (see Figure A-11). Another study using visuoconstructional and memory scores from the Rey-Osterrieth Complex Figure Test in children 6–11 years old observed significantly lower scores with increasing concurrent child single spot urinary fluoride even after adjusting for age (partial correlation coefficients, per log-mg/L increase = −0.29 and −0.27 for copy [p-value <0.001] and immediate recall [p-value <0.001], respectively [CIs not reported]) (Rocha-Amador et al. 2009). Although these children were also exposed to arsenic, the presence of arsenic could not explain the changes because, in the area with natural contamination by fluoride and arsenic (F–As), the test scores were not significantly associated with urinary arsenic levels (partial correlation coefficients, per log-mg/L increase = −0.05 and 0.02 for copy and immediate recall, respectively [CIs not reported]). The test scores were only marginally increased from fluoride alone when both fluoride and arsenic were included simultaneously in the model (partial correlation coefficients, per log-mg/L increase = −0.32 and −0.34 for copy and immediate recall, respectively [CIs not reported]) (Rocha-Amador et al. 2009) (see Figure A-10).

Commented [A82]: This sentence reflects revisions to characterize the test more precisely as the visuoconstructional and memory score from the Rey-Osterrieth Complex Figure Test, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 19): Please clarify what a construction task is (page 56). Do you mean a fine motor copy task?

Commented [A83]: The rest of this paragraph starting here reflects revisions to mention the results adjusted for both fluoride and arsenic, in response to the [redacted] Reviewer comment below; see DocH_Monograph for detailed response:

Reviewer comment (DocH_Monograph, page 19): Also on page 56 and highlighted in blue: this is unclear. Even though urinary arsenic is not associated with scores on these tasks, it could still very well be a confounder of the relationships between fluoride and the test scores.

Attention-related Disorders Including ADHD and Learning Disabilities in Children

Four studies evaluated attention-related disorders or learning disabilities (Barberio et al. 2017b; Bashash et al. 2018; Riddell et al. 2019; Wang et al. 2020a). All four studies found an association between increased fluoride and increased ADHD or learning disability; however, studies varied in the exposure metrics and outcomes measure. 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 significantly associated with a 2.84-point increase [95% CI: 0.84, 4.84; p-value = 0.0054] in the DSM-IV Inattention Index and a 2.54-point increase [95% CI: 0.44, 4.63; p-value = 0.0178] 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 significantly 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; p-value = 0.0176] in the DSM-IV ADHD Total Index and a 2.47-point increase [95% CI: 0.43, 4.50; p-value = 0.0175] in the ADHD Index) (see Figure A-11). 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 Conners' 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 Conners' Parent Rating Scale (Chinese version) to assess behavioral outcomes in children ages 7–13 years but found only a significant association between spot urinary fluoride concentrations in children (model adjusted for creatinine) and psychosomatic problems (adjusted OR for T-score >70 per 1-mg/L increase = 1.97 [95% CI: 1.19, 3.27; p-value = 0.009] and adjusted β per 1-mg/L increase = 4.01 [95% CI: 2.74, 5.28; p-value <0.001]). No associations were found between spot urinary fluoride and the ADHD index or other behavioral measures.

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 per 1-µmol/L increase = 1.02; 95% CI: 1.00, 1.03; p-value <0.05) (see Figure A-12); 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; p-value <0.05) and community water fluoridation status (adjusted OR per 1-mg/L increase = 1.21; 95% CI: 1.03, 1.42; p-value <0.05). A similar increase in the 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; p-value <0.05) and community fluoridation status (adjusted β per 1-mg/L increase = 0.11; 95% CI: 0.02, 0.20; p-value <0.05). As was observed with Barberio et al. (2017b), Riddell et al. (2019) did not observe associations between specific gravity-adjusted spot urinary fluoride concentrations and either ADHD diagnosis (adjusted OR per 1-mg/L increase = 0.96; 95% CI: 0.63, 1.46) or hyperactivity-inattention symptoms (adjusted β per 1-mg/L increase = 0.31; 95% CI: −0.04, 0.66).

Summary of Key Findings for Low Risk-of-bias Studies of Other Neurodevelopmental and Cognitive Effects in Children

In summary, the high-quality studies (i.e., studies with low potential for bias) provide evidence of an association between higher fluoride exposure and neurodevelopmental and cognitive effects in children other than IQ; however, the body of evidence is limited by heterogeneity in the outcomes evaluated and few directly comparable studies. Across these outcomes, eight of nine studies reported a significant association between fluoride exposure and a measure of neurodevelopment or cognition other than IQ, which provides support for the consistency in evidence based on children's IQ studies of an association between fluoride exposure and adverse effects on cognitive neurodevelopment.

Commented [A84]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer Comment (DocF_Monograph, pages 5 - 11) : High?

High Risk-of-bias Studies

High risk-of-bias studies (n = 6) also provide some evidence of associations between fluoride exposure and neurodevelopmental or cognitive effects in children other than effects on IQ, but the results are inconsistent and address different outcomes (Jin et al. 2016; Li et al. 1994

[translated in Li et al. 2008b]; Malin and Till 2015; Morgan et al. 1998; Mustafa et al. 2018; Shannon et al. 1986).

Risk of Bias for Neurodevelopmental or Cognitive Effect Studies in Children

The confidence in the human body of evidence was based on studies with the lowest potential for bias (i.e., studies that rated probably low or definitely low risk of bias for at least two of the three key risk-of-bias questions and did not have any other risk-of-bias concerns that would indicate serious issues with the studies). Each of the nine low risk-of-bias studies on other neurodevelopmental effects in children had little or no risk-of-bias concerns. Four of the nine studies were rated definitely low or probably low risk of bias for all risk-of-bias questions, and the remaining five studies were rated probably high risk of bias for a single question that was judged to have minimal impact on overall potential bias. None of the nine studies had a rating of definitely high risk of bias for any question. Although the nine low risk-of-bias studies had minimal or no concerns, the six studies with high overall potential for bias had several risk-of-bias concerns related to one or more of the three key risk-of-bias questions (confounding, exposure characterization, and outcome assessment). The key risk-of-bias questions are discussed below. Risk-of-bias ratings for other neurodevelopmental effect studies in children are available in Figure D-9 through Figure D-12 and Appendix E for the low and high risk-of-bias studies.

Confounding for Other Neurodevelopmental Studies in Children

Low Risk-of-bias Studies

As discussed above, there are nine studies considered to have low risk of bias when assessed across all risk-of-bias domains. Seven of nine low risk-of-bias studies were considered to have low potential for bias due to confounding because the authors addressed the three key covariates for all studies (age, sex, and socioeconomic status) and also addressed arsenic as a potential co-exposure of concern through study design or analysis. Other important covariates, including health factors, smoking, and parental characteristics, were also addressed in many of the low risk-of-bias studies. One of the studies (Bashash et al. 2018) examined several covariates 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 none of the sensitivity analyses indicated appreciable changes in the fluoride-related association with behaviors related to ADHD, nor was there evidence of effect modification between maternal urinary fluoride and sex.

Among the nine low risk-of-bias studies, two studies were identified that have potential for bias due to confounding (Rocha-Amador et al. 2009; Valdez Jimenez et al. 2017). Although both of these studies adjusted for several covariates through analysis or study design, Valdez Jimenez et al. (2017) did not address a potential concern for co-exposure to arsenic, and Rocha-Amador et al. (2009) does not appear to adjust for SES or address why it would not be a concern in the study population (see Appendix E for further details). Although these two studies have some potential for bias due to confounding, they are considered to have low potential for bias overall because they have low potential for bias for the other two key risk-of-bias questions (exposure characterization and outcome assessment), and no other major concerns for bias were identified.

Consistent with the IQ studies, bias due to confounding is not likely a concern for the low risk-of-bias studies.

High Risk-of-bias Studies

The six high risk-of-bias studies in the human body of evidence did not adequately address important covariates through study design or analysis. The same concerns due to potential confounding noted previously for the high risk-of-bias children's IQ studies were also present in the other neurodevelopmental high risk-of-bias studies, including not addressing the three key covariates for all studies (age, sex, SES) and/or not addressing potential co-exposures (e.g., arsenic) in areas of potential concern.

Exposure Characterization in Other Neurodevelopmental Studies in Children

Low Risk-of-bias Studies

There were no risk-of-bias concerns regarding exposure assessment in the low risk-of-bias studies. All of the low risk-of-bias studies had individual exposure data based on urine or water measures with appropriate analyses, and most of the urinary fluoride studies accounted for urinary dilution when appropriate. Although there are concerns related to the timing of urine samples (see the Risk-of-bias Considerations for Human Studies section for details), the studies that used maternal urine measured urinary fluoride multiple times throughout pregnancy (Bashash et al. 2017; Bashash et al. 2018; Valdez Jimenez et al. 2017). Another study demonstrated correlations between urinary fluoride and fluoride in the drinking water, fluorosis, or estimated dose based on water (Choi et al. 2015). Bashash et al. (2017) excluded exposure measurement outliers but found that doing so did not change the results in a meaningful way.

High Risk-of-bias Studies

A frequent critical limitation among the high risk-of-bias studies was lack of information regarding exposure or poor exposure characterization. In the high risk-of-bias studies that assessed the association between fluoride exposure and other neurodevelopmental and cognitive effects in children, fluoride exposure assessment was based on dental fluorosis, municipality-level water fluoridation prevalence data, number of years living in an area with fluorinated water, or group-level water samples. See the Exposure Characterization in IQ Studies section for further discussion on the limitations of exposure assessments in high risk-of-bias studies.

Outcome Assessment in Other Neurodevelopmental Studies in Children

Low Risk-of-bias Studies

The low risk-of-bias studies have few concerns regarding outcome assessment. Seven of the nine studies [i.e., all low risk-of-bias studies except Barberio et al. (2017b) and Riddell et al. (2019)] used appropriate methods for measuring other neurodevelopmental effects in the study population, and blinding of outcome assessors was either reported or not a concern in eight of the nine studies [i.e., all with the exception of Wang et al. (2020a)].

Among the nine low risk-of-bias studies, three were identified that have a potential for bias due to outcome assessment. One of the studies (Wang et al. 2020a) had potential concern for bias due to lack of information regarding the blinding of outcome assessors. Two of the studies (Barberio et al. 2017b; Riddell et al. 2019) were based on the same study population in Canada, where different questions were asked in Cycles 2 (2009–2011) and 3 (2012–2013) of the Canadian

Health Measures Survey (CHMS) to ascertain learning disabilities including ADHD. In Cycle 2, subjects were asked whether they had a learning disability diagnosed by a health professional and, if yes, were asked what kind. In Cycle 3, CHMS did not ask what kind of learning disability was diagnosed nor was a reason for the question omission provided. Because no reason was provided for the removal of the question, and because a question on learning disability without the specific diagnosis may be more prone to bias, this change in questioning from Cycles 2 to 3 is a potential concern. Blinding was not considered an issue in these two studies, but the methods for obtaining the information are considered to be less than ideal for measuring learning disabilities including ADHD. Although the questionnaire asked about a doctor's diagnosis of a learning disability, there was no confirmation with medical records. Moreover, these questionnaires were not validated like Conners' Rating Scales, which would have been a better method for assessing ADHD. Although the outcome assessment methods are less than ideal, there was no direct evidence that they were conducted incorrectly or that the methods would have biased the results in any specific direction. Because this was the only concern in these studies, they were considered to have low risk of bias overall.

High Risk-of-bias Studies

Among the studies on other neurodevelopmental effects with high potential for bias, there were several reasons for studies to be considered probably or definitely high risk of bias for outcome assessment. One study (Shannon et al. 1986) was considered to have probably high risk of bias based on lack of information regarding blinding of outcome assessors. One study was considered definitely high risk of bias because outcome was assessed based on a parent-completed questionnaire, and the study authors noted that the parents were informed of the study's intent and were requested to provide information on fluoride history. Other studies used outcome assessment methods that were not validated or utilized group-level measurements (i.e., school performance, working memory scores).

Confidence Assessment of Findings on Other Neurodevelopmental Effects in Children

The high-quality studies (i.e., studies with low potential for bias) provide some evidence of an association between fluoride exposure and other cognitive neurodevelopmental effects, including lower neurobehavioral scores in infants, cognitive effects other than IQ in children, and increased attention-related disorders including ADHD in children. However, due to limitations in the data set, including the heterogeneity in the outcomes assessed, a limited number of directly comparable studies, and differences in outcome assessment methods even when studies evaluated similar outcomes, there is low confidence based on this body of evidence that fluoride exposure is associated with other cognitive neurodevelopmental effects in children. Due to these limitations, the confidence assessment is not described in the same manner as the IQ in Children section or as outlined in Figure 1. Although there are limitations in the body of evidence, the low risk-of-bias studies demonstrate a relationship between higher fluoride exposure and neurodevelopmental effects, even in very young children, which supports the consistency in evidence shown in children's IQ studies of an association between fluoride exposure and adverse effects on cognitive neurodevelopment.

Commented [A85]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 19): This [second] sentence directly contradicts the leading, first sentence in this paragraph. The first sentence by itself is misleading. Note: reviewer struck out "evidence" and inserted "assumption" as follows […*provide ~~evidence~~ assumption of an association between fluoride exposure and other cognitive neurodevelopmental effects"*] with red text indicating editorial change.

Commented [A86]: The following two sentences reflect revisions to further clarify why we chose not to describe the confidence assessment in the same format and level of detail as in the IQ in Children section, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 21): Sufficiently supported, but it's unclear why the same format used for the IQ studies (pages 48-49) is not used here (page 59).

Cognitive Effects in Adults

Low Risk-of-bias Studies

Overview of Studies

Two low risk-of-bias cross-sectional studies evaluated the association between fluoride exposure and cognitive effect in adults (Jacqmin et al. 1994; Li et al. 2016). These two studies used the same test for cognitive function (i.e., Mini-Mental State or MMS Examination) and used drinking water fluoride levels to assess fluoride exposure. Li et al. (2016) also measured urinary fluoride. Both studies were cross-sectional in design. One was conducted in France (Jacqmin et al. 1994) and the other in China (Li et al. 2016). Both studies were conducted in older populations (i.e., over 60 or 65 years of age).

Table 8 provides a summary of study characteristics and key findings related to fluoride exposure and cognitive effects in adults for the two low risk-of-bias studies. The purpose of the table is to summarize key findings (independent of whether an association was found) from each study and is not meant to be a comprehensive summary of all results. For each study, results are summarized for each exposure measure assessed. Results from multiple analyses using the same exposure measure may not all be presented unless conflicting results were reported.

Commented [A87]: This sentence reflects revisions to include information on counts of studies per study design, in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 12): In general, it would be helpful to the reader to describe the included study designs earlier in the respective results sections (i.e., along with the total #s of included articles, describing the # of cross-sectional, prospective cohort, etc is important).

Table 8. Studies on Cognitive Function in Adults[a]

Study	Study Design (Location/Subjects) [n]	Exposure Measures and Summary Statistics	Assessment Timing	Outcome and Methods	Neurobehavioral Outcome Summary[b]
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: MMS Examination	No significant increase in the prevalence of cognitive impairment with increasing fluoride quartiles No statistical adjustment for covariates for prevalence rates
Li et al. (2016)	Cross-sectional China (Inner Mongolia)/adults [511]	Drinking water daily fluoride intake Mean (SD): 2.23 (2.23) (normal group), 3.62 (6.71) (cognitive impairment group) mg Urine Mean (SD): 1.46 (1.04) (normal group), 2.47 (2.88) (cognitive impairment group) mg/L Fluorosis score Mean (SD): 0.74 (0.98) (normal group), 1.29 (1.01) (cognitive impairment group)	Adults (ages ≥60 years)	Cognitive function: MMS Examination	Subjects with cognitive impairment had a significantly higher skeletal fluorosis score and urinary fluoride concentrations; odds of increasing severity of cognitive impairment increased with urinary fluoride concentrations but were not statistically significant; no significant association with total daily water fluoride intake 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

GM = geometric mean; MMS = Mini-Mental State.
[a]Includes low risk-of-bias studies.
[b]Associations between cognitive effects in adults and fluoride levels were reported quantitatively, when possible. For studies with multiple analyses and results, the table summarizes key findings and is not a comprehensive summary of all findings. Results also indicate when a study reported no association, provided as a qualitative statement of no association.

Summary of Results

Results from two low risk-of-bias studies in adults did not provide enough evidence to evaluate consistency when assessing evidence for a potential association between fluoride exposure and cognitive impairment (based on the MMS Examination) (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 A-13). In contrast, 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 in an analysis of 38 cognitively impaired cases and 38 controls matched for several covariates, including age, sex, education, alcohol consumption, and smoking (p-value <0.05). However, the authors found no significant association between cognitive impairment and total daily water fluoride intake (adjusted ORs per 1-mg/day increase = 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 per 1-mgL increase = 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).

Commented [A88]: This sentence reflects revisions to more accurately make a statement regarding the consistency in results across the two studies, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 22): The analysis between the two studies may be too simplistic. The French study was done in adults with not very high exposures to fluoride. In contrast the Chinese study compared adults with skeletal fluorosis (suggesting very high exposure) with others. It may be inaccurate to suggest that these two studies were not consistent. They may (consistently) show that relatively low exposures (even if above recommended) are not associated with cognitive outcomes, but very high exposures are.
This gets at [redacted] comments before about a lack of analysis regarding doses, dose effects, or thresholds.

High Risk-of-bias Studies

The results from five out of eight high risk-of-bias studies provide evidence of cognitive impairment in adults associated with higher levels of exposure to fluoride; however, there was heterogeneity in the outcomes assessed, a limited number of directly comparable studies, and some variability in results (e.g., variation in IQ results across studies). Due to the limited number of low risk-of-bias studies identified that assess cognitive impairment in adults, the results from the high risk-of-bias studies are summarized in greater detail below than had been done in this document for bodies of evidence for IQ in children and other neurodevelopmental and cognitive effects in children.

Commented [A89]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer comment (DocF_Monograph, page 9): High levels of

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 5 years of age, based on whether 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 38 years of age (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 on whether fluoride was able to reduce the risk associated with aluminum by competing with aluminum and reducing its bioavailability. Therefore, the study was considered inadequate to evaluate the association between fluoride and dementia (Still and Kelley 1980). A more recent study in Scotland evaluated dementia rates associated with aluminum and fluoride drinking water concentrations and observed a significant increased risk of

dementia per standard deviation increase in fluoride (p-value <0.001) with the risk of dementia more than double in the highest quartile of fluoride exposure (56.3 µg/L) compared to the lowest quartile (<44.4 µg/L). The authors also found a significantly increased risk of dementia associated with increased aluminum levels at all quartiles compared with the reference group (p-values <0.05) but found no statistical interaction between aluminum and fluoride levels in relation to dementia (Russ et al. 2019). Conversely, a study in China did not find a significant association between fluoride concentrations in the drinking water and risk for dementia (Liang et al. 2003). In addition to studies that reported on cognitive impairment and exposure to fluoride, two high risk-of-bias studies were identified that reported impaired motor and sensory function (Rotton et al. 1982) and a higher prevalence of self-reported headaches, insomnia, and lethargy (Sharma et al. 2009) associated with fluoride exposure.

Risk of Bias for Cognitive Effect Studies in Adults

Due to the small number of studies with a low potential for bias (see Figure D-13 and Figure D-14), the key risk-of-bias domains (confounding, exposure characterization, outcome assessment) are not discussed separately in respective subsections, as was done for the IQ in Children and Other Neurodevelopmental and Cognitive Effects in Children bodies of evidence. The high risk-of-bias studies had concerns across several domains (see Figure D-15 and Figure D-16), but there were still relatively few studies. Therefore, the discussion for high risk-of-bias studies is also not separated into subsections by key domain.

Low Risk-of-bias Studies

Both low risk-of-bias studies on cognitive effects in adults had little or no risk-of-bias concerns. One study was rated definitely low or probably low risk of bias for all risk-of-bias questions (Li et al. 2016), and the other study was rated probably high risk of bias for a single question that was judged to have minimal impact on overall potential bias (Jacqmin et al. 1994). Jacqmin et al. (1994) had potential concern for bias due to confounding because smoking was not addressed, which has the potential to impact risk for Alzheimer's disease and rates could vary by parish (the target population consisted of men and women from 75 civil parishes in southwestern France).

High Risk-of-bias Studies

There were several issues in the eight studies in adults considered to have high potential for bias. Four of the eight studies had potential concern for bias due to lack of information on the comparison groups, or the comparison groups were considered inappropriate. All eight studies had potential concern for bias regarding covariates not being addressed, including possible co-exposures in occupational studies (e.g., aluminum) and smoking. Five of the eight studies had potential concern for bias due to lack of information regarding exposure characterization or poor exposure characterization with the most utilized exposure measure in these studies being a comparison between exposed and unexposed areas. In one case (Broadbent et al. 2015), multiple sources of fluoride exposure were assessed separately without properly controlling for the other sources of exposure, which could bias the results (see Exposure Characterization in IQ Studies for further details). Five studies also had potential for bias based on limitations in the outcome assessment, which was mainly due to lack of blinding of outcome assessors, lack of validation of the methods, or lack of sufficient details on how the outcomes were assessed.

Confidence Assessment of Findings on Cognitive Effects in Adults

The body of evidence available to examine the association between exposure to fluoride and cognitive effects in adults is limited to two low risk-of-bias cross-sectional studies. Due to the limited number of studies and a lack of evidence of an effect, there is low confidence based on this body of evidence that fluoride exposure is associated with cognitive effects in adults.

Mechanistic Data in Humans

Eight low risk-of-bias studies that evaluated fluoride exposure and mechanistic data in humans were considered potentially relevant to neurological effects. Effects on the thyroid were specifically evaluated because the NRC 2006 report identified this as a possible effect of fluoride (NRC 2006), and changes in thyroid hormones have been identified as a mechanism for neurodevelopmental effects (Haschek and Rousseaux 1991). These included effects on thyroid hormones in children (Kheradpisheh et al. 2018a; Kheradpisheh et al. 2018b; Malin et al. 2018), adults (Kheradpisheh et al. 2018a; Kheradpisheh et al. 2018b; 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 D-17 and Figure D-18). Although the low 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 low risk-of-bias studies that reported on changes in thyroid hormones, three studies were conducted in children (Kumar et al. 2018; Singh et al. 2014; Zhang et al. 2015b) 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), whereas 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). When all children (with and without dental fluorosis) in the endemic area were compared with children from the non-endemic area, the TSH levels were higher in children from the fluorosis-endemic area, although results did not reach statistical significance ($p = 0.057$). 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 associations between fluoride and 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 low 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).

Sixteen high risk-of-bias studies were available that evaluated mechanistic data in humans associated with fluoride exposure, including effects on thyroid hormones in children (n = 9 studies), thyroid hormones in adults (Michael et al. 1996; Yasmin et al. 2013), catecholamines in adults (Michael et al. 1996) or in subjects of unknown ages (Chinoy and Narayana 1992), acetylcholinesterase (AChE) or serotonin levels in children (Lu et al. 2019; Pratap et al. 2013), 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 low risk-of-bias studies, the high 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 high risk-of-bias studies (see Figure D-19 and Figure D-20), varying results were reported in 11 studies that evaluated associations between fluoride exposure and thyroid hormones, and a few of these studies (Lin et al. 1991; Wang et al. 2001; Yang et al. 1994 [translated in Yang et al. 2008]) 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 low 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 association. Six of the nine high 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 high 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 associations between fluoride exposure and 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 low and high risk-of-bias studies that evaluated associations between fluoride exposure and 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

Commented [A90]: Editorial change made during formatting

displayed variation by age in the associations between fluoride and 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.

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**

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

Interactive figure and additional study details are available at https://public.tableau.com/app/profile/ntp.visuals/viz/FluorideTableauDashboards/ReadMe. This figure displays study counts for low risk-of-bias studies in both children and adults, as these counts are most relevant to the summary of fluoride-related mechanistic effects in low risk-of-bias studies. Counts for high risk-of bias studies and studies by age (i.e., children, adults, or children/adults combined) can also be accessed in the interactive figure. Study counts are tabulated by significance (unless study footnotes in the interactive figure 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.

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**

Figure 8. Number of High Risk-of-bias Studies that Evaluated Thyroid Hormones in Children by Endpoint and Direction of Association

Interactive figure and additional study details are available at https://public.tableau.com/app/profile/ntp.visuals/viz/FluorideTableauDashboards/ReadMe. This figure displays study counts for high risk-of-bias studies in children, as these counts are most relevant to the summary of associations between fluoride and thyroid hormones in high risk-of-bias studies. Counts for low risk-of bias studies, studies in adults, or all studies combined can also be accessed in the interactive figure. Study counts are tabulated by significance (unless study footnotes in the interactive figure 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.

In addition to evaluating thyroid hormone levels, a few high 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 (it was not reported whether subjects were children or adults) compared with a non-endemic region (Chinoy and Narayana 1992). A separate study reported that Sserum epinephrine and norepinephrine (referred to as adrenaline and noradrenaline in the study) were significantly increased in adults in a fluoride-endemic area (fluoride in the drinking water ranged from 1.0–6.53 ppm) compared with 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 with a lower fluoride region (Pratap et al. 2013). Serum serotonin was significantly increased in children from Turkey who were drinking water containing 2.5 mg/L of fluoride compared with 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 with a control area (Du et al. 1992 [translated in Du et al. 2008]; Yu et al. 1996 [translated in Yu et al. 2008]).

Commented [A91]: Change made in response to the [redacted] comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 19): For the non-health professional reader, the use of different nomenclature for the same neurotransmitter is confusing.

There are also two more recent low 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 association between log urinary 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).

Animal Learning and Memory Data

NTP provided a review of the experimental animal evidence in the earlier draft monographs (NTP 2020) and agrees with the NASEM committee's comments (NASEM 2020; 2021) that the experimental animal database is of poor quality, with many studies suffering from major reporting deficiencies (*Sup01_Monograph_NASEM_Feb_2021.pdf*). NTP 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. Overall, these general issues and deficiencies with the experimental animal database led to NTP's conclusion that the animal studies are currently *inadequate* to inform the question of an association between fluoride exposure and neurodevelopmental and cognitive effects in humans. Therefore, this systematic review does not include an experimental animal section.

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 Appendix F); however, the mechanisms underlying fluoride-associated cognitive neurodevelopmental effects are not well characterized, and review of the data did not identify a mode of action for fluoride effects on IQ in children. 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 backcalculated into equivalent drinking water concentrations for comparison) still provided a sufficient number of studies for evaluation of these mechanistic endpoints. This evaluation is provided in Appendix F. Neurotransmitter and biochemical changes in the brain and neurons were considered 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 Appendix F). 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.

In Vitro Data on Neurodevelopmental or Cognitive Effects

Although in vitro studies were identified as part of the systematic review process, NTP determined that the information on neurological effects 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.

Discussion

This systematic review evaluated the available animal and human literature concerning the association between fluoride exposure and cognitive neurodevelopment. The available data on potential mechanisms to evaluate biological plausibility were also assessed. The ~~potential~~ health benefits of fluoride with respect to oral health are acknowledged but are not the focus of this review.

This review extended NTP's previous evaluation of the experimental animal data (NTP 2016). Although the animal data provide some evidence of effects of fluoride on neurodevelopment, they give little insight into the question of whether fluoride influences IQ. This is due to deficiencies identified in the animal body of evidence. Mechanistic studies in humans provide some evidence of adverse neurological effects of fluoride. However, these studies were too heterogenous and limited in number to make any determination on biological plausibility.

The literature on adults is also limited; therefore, it was determined that there is low confidence in the body of evidence from studies that evaluate fluoride exposure and adult cognition. Compared to the literature in adults, there is a much more extensive literature in children.

The literature in children was separated into studies assessing IQ and studies assessing other cognitive or neurodevelopmental outcomes. There is low confidence in the body of evidence from studies that evaluate fluoride exposure and other cognitive or neurodevelopmental outcomes in children. This body of evidence is made up of~~Altogether, the results from eight of~~ nine high-quality studies (three prospective cohort and ~~five~~ six cross-sectional studies from seven different study populations) and six low-quality studies. Eight of the nine high-quality studies provide some evidence that fluoride is associated with other cognitive or neurodevelopmental outcomes in children. The data also suggest that neurodevelopmental effects occur in very young children. However, the confidence in this body of evidence is low because the number of studies is limited, and there is too much heterogeneity in the outcomes measured, ages assessed, and methods used~~,~~ to directly compare studies of any one outcome. Additional studies on outcomes such as attention-deficit hyperactivity disorder (ADHD) and other attention-related disorders, where there is some evidence of an effect of fluoride exposure, would be necessary to critically assess the data.

Most of the epidemiological studies (n = 72) assessed the association between fluoride exposure and IQ in children. Although all studies, both high- and low-quality, were considered, this evaluation focuses on the high-quality, low risk-of-bias studies in children for two reasons. First, there are fewer limitations and greater confidence in the results of the high-quality studies. Second, there is a relatively large number of high-quality studies (n = 19), such that the body of evidence from these studies could be used to evaluate confidence in the association between fluoride exposure and changes in children's IQ.

This review finds, with moderate confidence, that higher fluoride exposure is associated with lower IQ in children. The association between higher fluoride exposure and lower IQ in children was consistent across different study populations, study locations, study quality/risk-of-bias determinations, study designs, exposure measures, and types of exposure data (group-level and individual-level). There were 19 low risk-of-bias studies that were conducted in 15 study populations, across 5 countries, and evaluating more than 7,000 children. Of these 19 studies, 18

Commented [A92]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, page 2): Please confirm with [redacted] for references, but [redacted] believe there is a significant amount of data out there to suggest that it's more than just "potential health benefits." Recommend expanding upon this a bit more to describe some of the health benefits that have been shown.

Commented [A93]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 20): This appears contrary to the preceeding sentence and is not valuable to present description of the details of low-confidence studies.

Commented [A94]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to several related reviewer comments:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

reported an association between higher fluoride exposure [e.g., represented by populations whose total fluoride exposure approximated or exceeded the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] and lower IQ. These include 3 prospective cohort studies and 15 cross-sectional studies (12 of which indicated that exposure likely preceded the outcome). Forty-six of 53 low-quality studies in children also reported an association between higher fluoride exposure and lower IQ.

Many studies in this assessment relied on drinking-water fluoride levels (both group-level measures and individual-level measures), rather than measures of total fluoride exposure, to establish exposed versus "unexposed" or reference groups. Although fluoride in water is a major source of exposure [comprising 40% to 70% of total exposure (US EPA 2010)], other sources of fluoride provide variable amounts that depend on personal preferences and habits. The use of dental products containing fluoride and consuming foods and beverages prepared with fluoridated water can also result in measurable exposures (US EPA 2010). Green et al. (2019) suggested that significant exposures occur from black tea consumption. Thus, drinking water fluoride levels may, but usually do not, reflect total fluoride exposure. This could be a potential limitation in studies that rely on water fluoride data to assess fluoride exposure (in particular, earlier studies). However, because water is only part of a person's total exposure to fluoride, this limitation would likely result in an underestimate of exposure to fluoride. In addition, this limitation is less of a concern in areas where fluoride in the drinking water is high because drinking water likely contributes a large proportion of the total fluoride intake in those areas as compared with areas where fluoride in the drinking water is lower.

This review found that the quality of exposure assessment has improved over the years. More recent studies by Valdez Jimenez et al. (2017), Bashash et al. (2017), and Green et al. (2019) used individual measures of urinary fluoride, either maternal urine collected prenatally or children's urine, which ~~confirmed~~ support the association between higher total fluoride exposure and lower children's IQ and other cognitive neurodevelopmental effects. Studies using different types of exposure measures reported similar findings of an association (NRC 2006), which strengthens confidence in earlier studies that reported IQ deficits with ~~increasing~~ high group-level fluoride exposure. However, there is less certainty in the quantitative estimates of the magnitude of IQ deficits from earlier studies that used group-level exposure measures than the estimates from more recent studies that used individual-level exposure measures.

It is worth noting that there are circumstances wherein typical children's water consumption considered with water fluoride levels may substantially underestimate total fluoride exposure. One example is bottle-fed infants wherein nutrition is provided by powdered formula that is rehydrated with fluoridated water (Till et al. 2020). To decrease an exclusively formula-fed infant's exposure to fluoride, for the purpose of reducing risk of dental fluorosis, the Centers for Disease Control and Prevention recommends using low-fluoride bottled water to mix with infant formula (CDC 2015). A few studies also support the possibility of heightened sensitivities to the detrimental cognitive effects of fluoride exposure in individuals with certain genetic polymorphisms in dopamine receptor D2 or catechol-O-methyltransferase (Cui et al. 2018; Zhang et al. 2015b), potentially impacting dopamine catabolism and receptor sensitivity. Differential exposures to fluoride and genetic susceptibilities of children to fluoride may represent special situations that would appear to warrant further research.

Commented [A95]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Review comment (DocC_Monograph, page22): The association is not causal and therefore, suggest changing the word [confirmed] to "support"... the association.

Commented [A96]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 22): This statement appears to imply that there are no studies with negative findings and the sentence asserts a dose-response relationship for which data are incomplete.

The following section briefly recaps the strength of the epidemiological evidence for an association between fluoride exposure and cognitive neurodevelopmental deficits. This is followed by a more detailed listing of limitations of the evidence base and limitations of the systematic review, with some suggestions of areas where further research may be most beneficial.

Limitations of the Evidence Base

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

- Few studies are available that assessed the association between fluoride exposure and cognitive function (particularly IQ) in adults and attention-related disorders including ADHD in children and adults.
- Heterogeneity in outcomes was assessed for other neurobehavioral outcomes, limiting the assessment of other possible effects in children.
- Studies rarely separated the results by sex or provided information to indicate that sex was not a modifying factor.
- Associations between lower total fluoride exposure [e.g., represented by populations whose total fluoride exposure was lower than the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] and children's IQ remain unclear. More studies at lower exposure levels are needed to fully understand potential associations in ranges typically found in the United States (i.e., <1.5 mg/L in water). However, it should be noted that, as of April 2020, CWS supplying water with ≥1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population (~1.9 million people) (CDC Division of Oral Health 2020). This indicates that the moderate confidence in the association between higher fluoride exposure and lower IQ is relevant, at a minimum to children living in these areas of the United States where fluoride in drinking water is known to be at or above 1.5 mg/L. This is only compounded by additional exposures to fluoride from other sources.
- No studies investigating the association between fluoride exposure and neurodevelopmental or cognitive effects in adults or children have been conducted in the United States.
- No studies are available to evaluate lifelong exposure in adults, or fluoride exposure over a child's lifetime and neurodevelopmental or cognitive changes over time.
- The database does not allow for comparison of ages and possible changes at different developmental stages in children to assess if there is a delay in development or if associations persist.
- The database does not allow for establishing clear correlations between prenatal and postnatal exposures.

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

- Many of the original publications were in a non-English language and provided limited details on methodology.

Commented [A97]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response:

Reviewer comment (DocF_Monograph, page 16): Elaborate a little further.

Commented [A98]: The next three limitations (starting with this one) were added in response to [redacted] comments from the [redacted] Reviewer listed below; see DocH_Monograph for detailed response. *Note a* [redacted] *Reviewer also suggested an issue that was addressed in one of these limitations and is described in the subsequent comment (see DocJ_Monograph for detailed response)*:

Reviewer comment (DocH_Monograph, page 7): Cumulative exposure: The authors should make clear that exposure during gestation likely implies that there is continuing exposure in the post natal period. Further, these two exposure periods are likely highly correlated, making conclusions regarding a critical period of exposure difficult. The converse is also true – i.e. if exposure is measured in the post natal or childhood period, and especially if it is from drinking water, then there was likely exposure in the prenatal period as well.

Reviewer comment (DocH_Monograph, page 14): The results also need to be interpreted based on age of test administration. Some higher order functions do not develop until later ages and thus cannot be tested well in younger children. Also, as with other neurotoxins, deficits can occur at a variety of ages, and either persist or not. So the age at assessment becomes an important variable in the interpretation of findings and should be accounted for in the discussion.

Commented [A99]: This limitation was added in response to the [redacted] Reviewer comment below; see DocJ_Monograph for detailed response:

Reviewer comment (DocJ_Monograph, page 22): Yes, the approach used to assess risk of bias was clearly described and generally appropriate.
Though, it is unclear whether these studies adequately captured a critical fluoride exposure window likely to impact neurocognitive health (i.e., does fluoride exposure in older adulthood impact neurocognitive health?) For example, lifelong fluoride exposure, and/or fluoride exposure at different lifestages that may be more critical to neurocognitive development, were not captured in these cross-sectional studies. Thus, it raises questions as to whether these cross-sectional studies are truly "low risk of bias," or are "lower" risk of bias than others.

- 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 in a few studies, 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 still may have 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.
- Studies did not provide sufficient direct information (e.g., participation rates or methods for selection) to evaluate selection bias.
- Failure to address important covariates was an issue for many studies. Some studies conducted simple statistical analyses without accounting for any covariates in the analysis, although many noted similarities between the study populations. In cases where adjustments in analyses were made, often these studies did not account for covariates considered critical for that study population and outcome including co-exposures.
- 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 in areas where these substances were likely to occur.
- 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 and mechanistic evidence base include:

- The overall quality of the experimental animal studies is poor, and there are relatively few well-designed and well-performed studies at lower fluoride exposure levels (i.e., <20 ppm, which is roughly equivalent to human exposure of <4 ppm).
- The understanding of the specific molecular events responsible for fluoride's adverse effects on neurobehavioral function is poor.

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.

Strengths of the Evidence Base

Strengths in the epidemiological evidence base include:

- There are 72 studies directly addressing the relationship between fluoride exposure and children's IQ.
- There are 12 high-quality cross-sectional studies with low risk of bias providing evidence that exposure occurred prior to outcome assessment in those studies.
- Studies are from diverse geographic locations that included data for more than 7,000 children.

- There are 19 high-quality studies evaluating the same outcome (i.e., IQ) and 9 evaluating other neurodevelopmental outcomes.
- Reported ~~responses~~ associations between ~~to~~higher fluoride exposure and lower children's IQ are consistent in the vast majority of studies of both low and high quality.
- Reported ~~responses~~ associations between ~~to~~higher fluoride exposure and lower children's IQ are consistent across different study populations, study designs, and exposure measures.
- Findings of studies with group- and individual-level information on exposure and outcomes are similar.
- A wide variety of important covariates are either addressed by study design or captured across the evidence base, with no consistent patterns that would suggest an alternative explanation.

Limitations of the Systematic Review

This systematic review has few limitations. The human body of evidence included a large database of observational studies. Most of the observational studies were cross-sectional; however, 12 of these were considered to provide sufficient evidence that exposure occurred prior to the outcome. 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 these studies did not include specific information on thyroid hormones that could 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 because 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.

The supplemental literature search for non-English-language studies not indexed in traditional databases supports the comprehensive nature of the literature search strategy for this systematic review. In the absence of guidance on the most complete non-English-language databases that may contain health studies of fluoride, a standard systematic review approach for database selection was followed whereby a set of exemplar documents, called 'seed studies' were used. Databases were selected that identified non-English-language studies of fluoride that we were aware of and were not captured in searches of databases from the main literature search. This informed approach influenced the selection process; however, this is not considered a limitation because it provided an objective measure by which to compare databases. Following the recommendation of the NASEM committee in its review of the September 16, 2020, draft monograph, the experimental animal section has been removed and is not included in this monograph. Although the deficiencies identified in the animal body of evidence support this removal (see Animal Learning and Memory Data for further explanation), NTP acknowledges

Commented [A100]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 23): Of the 72 studies, there are some with equivocal or contrary results.

Commented [A101]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to related reviewer comment:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High?

Commented [A102]: Change made in response to the [redacted] Reviewer comment below; see DocC_Monograph for detailed response:

Reviewer comment (DocC_Monograph, page 23): This statement is contrary to the earlier declarations of significant heterogeneity across studies and thus, high risk of bias.

Commented [A103]: The following four sentences were added in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 4): Potential bias as there was no systematic selection of Chinese databases to be searched. Two databases were selected because they contained studies of which the authors were aware.

Recommendation: As NASEM noted, this introduces potential bias. [redacted] suggests this be added as a limitation.

that the absence of the experimental animal data is a limitation of this systematic review. For the purpose of this review, NTP considers the experimental animal data to be *inadequate* to inform whether fluoride exposure is associated with cognitive effects (including cognitive neurodevelopmental effects) in humans.

Summary

This systematic review evaluated the available animal and human literature concerning the association between fluoride exposure and cognitive neurodevelopment. The available data on potential mechanisms to evaluate biological plausibility were also assessed. Existing animal studies provide little insight into the question of whether fluoride exposure affects IQ. Human mechanistic studies were too heterogenous and limited in number to make any determination on biological plausibility. The body of evidence from studies on adults is also limited and provides low confidence that fluoride exposure is associated with adverse effects on adult cognition. There is, however, a large body of evidence on IQ effects in children. There is also some evidence that higher fluoride exposure is associated with other neurodevelopmental and cognitive effects; although, because of the heterogeneity of the outcomes, there is low confidence in the literature for these other effects. This review finds, with moderate confidence, that higher fluoride exposure [e.g., represented by populations whose total fluoride exposure approximates or exceeds the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] is consistently associated with lower IQ in children. More studies are needed to fully understand the potential for lower fluoride exposure to affect children's IQ.

Commented [A104]: Change made in response to the [redacted] Reviewer comment below; see DocF_Monograph for detailed response to several related reviewer comments:

Reviewer comment (DocF_Monograph, pages 5 - 11) : High

Commented [A105]: The following two sentences reflect revisions in response to related comments from [redacted] Reviewers listed below to focus the emphasis on the data on which we base our confidence statement and to acknowledge the need for further studies at lower exposure levels; see DocG_Monograph and DocH_Monograph, respectively for detailed response:

Reviewer comment (DocG_Monograph, page 18): An extension of a prior comment, on page 48 (Confidence Assessment of Findings on IQ in Children), the review does not provide evidence to support the statement that the high fluoride exposure should be interpreted as "mainly greater than the WHO Drinking Water Quality Guideline [≥ 1.5 mg/L]".

Reviewer comment (DocH_Monograph, page15): In general, the confidence rating in the body of evidence for this outcome is supported. However, several concerns necessitate a refinement of this confidence rating.
In agree with the prior review in that conclusions can only be made above the WHO drinking water limit for fluoride. It seems as though there is a lack of dose response curve estimation for lower levels of exposure, so an inference cannot be made over the entire range of exposure. Indeed, it is this lower dose range that is of interest for the US population.

Commented [A106]: This sentence and similar sentences throughout the monograph were revised to add "higher" as a qualifier when appropriate to describe fluoride exposures and to provide a benchmark (the 1.5 mg F/L WHO Guidelines for Drinking-Water Quality) to aid in describing total exposures above which moderate confidence was determined for children's IQ studies, in response to the [redacted] Reviewer comment below; see DocG_Monograph for detailed response:

Reviewer comment (DocG_Monograph, page 14): Throughout, it is important to talk about *higher* exposure to fluoride, not just exposure to fluoride. Everyone is exposed to fluoride so describing the at-risk group as being exposed to fluoride is meaningless (and confusing).

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. https://doi.org/10.1139/cjpp-2016-0641

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.

Arbuckle TE, Fraser WD, Fisher M, Davis K, Liang CL, Lupien N, Bastien S, Velez MP, von Dadelszen P, Hemmings DG et al. 2013. Cohort profile: The maternal-infant research on environmental chemicals research platform. Paediatr Perinat Epidemiol. 27(4):415-425. https://doi.org/10.1111/ppe.12061

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.

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. https://doi.org/10.4103/pm.pm_378_17

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. https://doi.org/10.1136/jech-2017-209129

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. https://doi.org/10.17269/cjph.108.5951

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. https://doi.org/10.1016/j.reprotox.2018.07.078

Bashash M, Thomas D, Hu H, Martinez-Mier EA, Sanchez BN, Basu N, Peterson KE, Ettinger AS, Wright R, Zhang Z et al. 2017. Prenatal fluoride exposure and cognitive outcomes in

children at 4 and 6-12 years of age in Mexico. Environ Health Perspect. 125(9):1-12. https://doi.org/10.1289/ehp655

Bashash M, Marchand M, Hu H, Till C, Martinez-Mier EA, Sanchez BN, Basu N, Peterson KE, Green R, Schnaas L et al. 2018. Prenatal fluoride exposure and attention deficit hyperactivity disorder (ADHD) symptoms in children at 6-12 years of age in Mexico City. Environ Int. 121(Pt 1):658-666. https://doi.org/10.1016/j.envint.2018.09.017

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.

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

CDC Division of Oral Health. 2020. Personal communication. September 3, 2020.

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

Centers for Disease Control and Prevention (CDC). 2015. Community water fluoridation FAQs: Infant formula Atlanta, GA: Centers for Disease Control and Prevention. https://www.cdc.gov/fluoridation/faqs/infant-formula.html. [22 September 2021]

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.

Cui Y, Zhang B, Ma J, Wang Y, Zhao L, Hou C, Yu J, Zhao Y, Zhang Z, Nie J et al. 2018. Dopamine receptor D2 gene polymorphism, urine fluoride, and intelligence impairment of children in China: A school-based cross-sectional study. Ecotoxicol Environ Saf. 165:270-277. https://doi.org/10.1016/j.ecoenv.2018.09.018

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. https://doi.org/10.1016/j.neulet.2020.134981

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.

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. https://doi.org/10.1016/j.puhe.2017.08.013

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

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.

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.

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. https://doi.org/10.1177/0192623315596858

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.

Green R, Rubenstein J, Popoli R, Capulong R, Till C. 2020. Sex-specific neurotoxic effects of early-life exposure to fluoride: A review of the epidemiologic and animal literature. Current Epidemiology Reports. 7(4):263-273. https://doi.org/10.1007/s40471-020-00246-1

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 Health & 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 et al. 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 64(4):383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026

Haschek W, Rousseaux C. 1991. Handbook of toxicologic pathology. 1st ed.: Academic Press.

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. 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. New York, NY: John Wiley & Sons.

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 et al. 2020. SWIFT-Active Screener: Accelerated document screening

through active learning and integrated recall estimation. Environ Int. 138:105623. https://doi.org/10.1016/j.envint.2020.105623

Ibarluzea J, Gallastegi M, Santa-Marina L, Jiménez Zabala A, Arranz E, Molinuevo A, Lopez-Espinosa MJ, Ballester F, Villanueva CM, Riano I et al. 2021. Prenatal exposure to fluoride and neuropsychological development in early childhood: 1-to 4 years old children. Environ Res. 207:112181. https://doi.org/10.1016/j.envres.2021.112181

International Programme on Chemical Safety (IPCS). 2002. Fluorides. Geneva: World Health Organization, International Programme on Chemical Safety. Environmental Health Criteria 227. https://inchem.org/documents/ehc/ehc/ehc227.htm.

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 et al. 2019. Maternal supplementation of low dose fluoride alleviates adverse perinatal outcomes following exposure to intrauterine inflammation. Scientific Reports. 9(1):2575. https://doi.org/10.1038/s41598-018-38241-8

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

Jin T, Han T, Wei Y, Wu Y, Wang Z, Zhang H. 2016. Investigation on working memory level of children aged 8-12 years in coal-burning fluorosis area. Journal of environment and health.409-411.

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

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. https://doi.org/10.1080/01480545.2017.1321009

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. https://doi.org/10.1007/s10661-017-6288-5

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. https://doi.org/10.1007/s10661-018-6501-1

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. https://doi.org/10.1007/s40201-018-0290-x

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. https://doi.org/10.1038/s41598-018-20696-4

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.

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, Q.L. S, 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 et al. 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 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.

Liang C, Ji R, Cao J, Jiang Y, Yu B, Ma F, Wu Y, Ying B, Zhang Y, Sun S et al. 2003. Study on the relationship between drinking water trace elements and cognitive ability of the elderly. Health Res.436-440.

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 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 et al. 2019. Fluoride related changes in behavioral outcomes may relate to increased serotonin. Physiol Behav. 206:76-83. https://doi.org/10.1016/j.physbeh.2019.02.017

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. https://doi.org/10.1016/j.chemosphere.2017.08.166

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. https://doi.org/10.1016/j.envint.2018.09.026

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. https://doi.org/10.1007/s12640-018-9870-x

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; October 25 2016; 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. https://doi.org/10.1371/journal.pmed.1000097~~

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.

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. https://doi.org/10.13040/IJPSR.0975-8232.9(8).3247-56

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

National Academies of Sciences Engineering and Medicine (NASEM). 2021. Review of the revised NTP Monograph on the systematic review of fluoride exposure and neurodevelopmental and cognitive health effects: A letter report. Washington, DC: National Academies of Sciences, Engineering and Medicine. https://doi.org/10.17226/26030.

National Institute for Occupational Safety and Health (NIOSH). 1984. Fluoride in urine. Washington, DC: National Institute for Occupational Safety and Health. Method 8308.

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

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

National Toxicology Program (NTP). 2019. Health Assessment Workspace Collaborative (HAWC) Page: Flouride (2019). Research Triangle Park, NC: U.S. Department of Health and Human Services, National Institutes of Health, National Toxicology Program. https://hawcproject.org/assessment/405/.

National Toxicology Program (NTP). 2020. Revised draft NTP Monograph on the systematic review of fluoride exposure and neurodevelopmental and cognitive health effects. Research Triangle Park, NC: National Toxicology Program. https://www.nationalacademies.org/event/10-19-2020/docs/DDA97C9170D1A255D69C004CEB77B698E8D005011EFB.

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. https://doi.org/10.1016/j.envpol.2017.09.015

Nkpaa KW, Onyeso GI. 2018. Rutin attenuates neurobehavioral deficits, oxidative stress, neuro-inflammation and apoptosis in fluoride treated rats. Neurosci Lett. 682:92-99. https://doi.org/10.1016/j.neulet.2018.06.023

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE , Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. 2021. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. PLoS Med. 18(3):e1003583. 10.1371/journal.pmed.1003583

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. https://doi.org/10.1126/science.aba1510

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.

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. https://doi.org/10.1016/j.envint.2019.105190

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.

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.

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. https://doi.org/10.1192/bjp.2018.287

Santa-Marina L, Jimenez-Zabala A, Molinuevo A, Lopez-Espinosa M, Villanueva C, Riano I, Ballester F, Sunyer J, Tardon A, Ibarluzea J. 2019. Fluorinated water consumption in pregnancy and neuropsychological development of children at 14 months and 4 years of age. Environ Epidemiol. 3:386-387. https://dx.doi.org/10.1097/01.EE9.0000610304.33479.18

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.

Scientific Committee on Health and Environmental Risks (SCHER). 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. Scientific Committee on Health and Environmental Risks. http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_139.pdf. [19 August 2019]

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.

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

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.

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, 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 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, 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. https://doi.org/10.1289/ehp3546

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. https://doi.org/10.1016/j.envint.2019.105315

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.

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

U.S. Environmental Protection Agency (US EPA). 2010. Fluoride: Exposure and relative source contribution analysis. Washington, DC: U.S. Environmental Protection Agency. 820-R-10-015. http://www.epa.gov/dwstandardsregulations/fluoride-risk-assessment-and-relative-source-contribution. [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. https://doi.org/10.1016/j.neuro.2016.12.011

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. https://doi.org/10.1159/000279325

Wang A, Duan L, Huang H, Ma J, Zhang Y, Ma Q, Guo Y, Li Z, Cheng X, Zhu J et al. 2020a. Association between fluoride exposure and behavioural outcomes of school-age children: A pilot study in China. Int J Environ Health Res.1-10. https://doi.org/10.1080/09603123.2020.1747601

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 et al. 2020b. Thyroid function, intelligence, and low-moderate fluoride exposure among Chinese school-age children. Environ Int. 134:105229. https://doi.org/10.1016/j.envint.2019.105229

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.

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.

Waugh DT. 2019. Fluoride exposure induces inhibition of sodium/iodide symporter (NIS) contributing to impaired iodine absorption and iodine deficiency: Molecular mechanisms of inhibition and implications for public health. Int J Environ Res Public Health. 16(6). https://doi.org/10.3390/ijerph16061086

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

World Health Organization (WHO). 2017. Guidelines for drinking-water quality. World Health Organization. 4th ed. + 1st add. https://apps.who.int/iris/handle/10665/254637.

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, 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, 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.

Yang Y, Wang X, X. G, 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.

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 et al. 2018. Threshold effects of moderately excessive fluoride exposure on children's health: A potential association between dental fluorosis and loss of excellent intelligence. Environ Int. 118:116-124. https://doi.org/10.1016/j.envint.2018.05.042

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.

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 et al. 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 Q, Niu Q, Chen J, Xia T, Zhou G, Li P, Dong L, Xu C, Tian Z, Luo C et al. 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. https://doi.org/10.1007/s00204-019-02390-0

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. https://doi.org/10.7150/thno.42387

Zhou G, Tang S, Yang L, Niu Q, Chen J, Xia T, Wang S, Wang M, Zhao Q, Liu L et al. 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. https://doi.org/10.1016/j.taap.2019.114608

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 et al. 2012. Environmental fluoride exposure and reproductive hormones in male living in endemic fluorosis villages in China. Life Sci J. 9(4):1-7.

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. https://doi.org/10.1111/j.1600-0528.2006.00269.x

Appendix A. Data Figures: Neurodevelopmental or Cognitive Effects and Outcomes

Figures

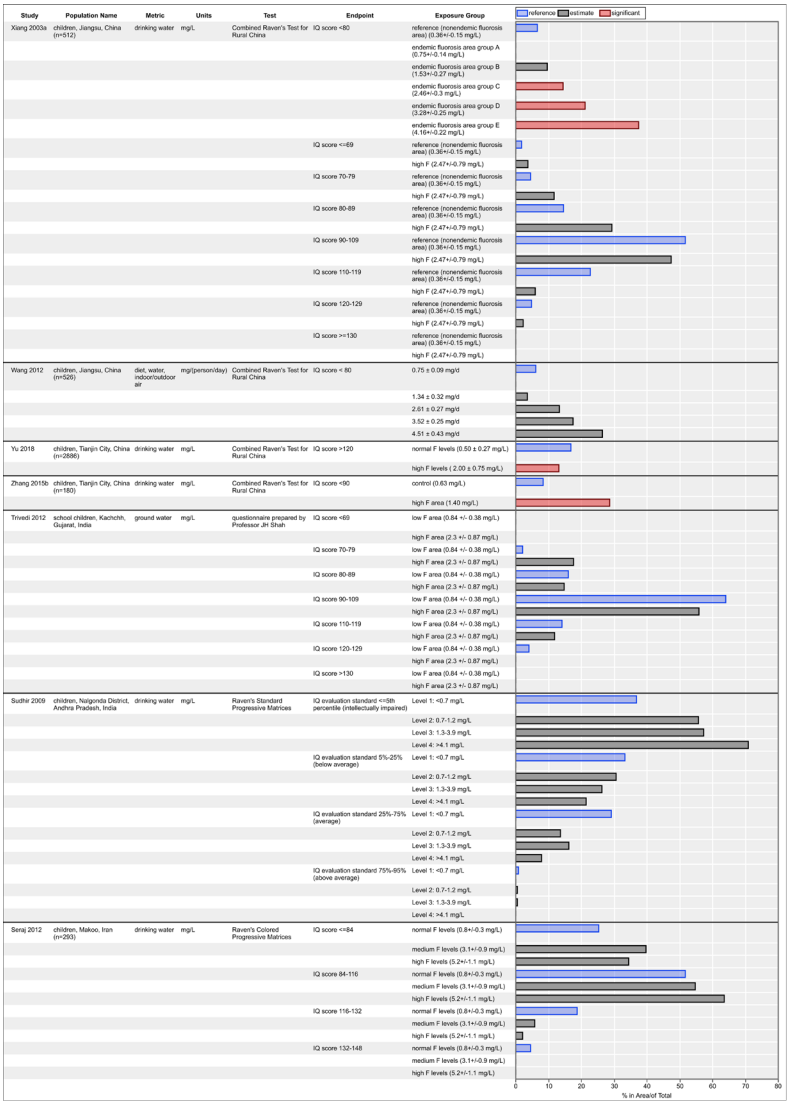


Figure A-1. Distribution of IQ in Children by Fluoride Exposure (Low Risk-of-bias Studies; Presented as % in Area or % of Total Group)

Reference group indicated by blue bars; other bars represent response estimates with red indicating statistical significance compared with the reference group.
An interactive version of Figure A-1 and additional study details in HAWC here (NTP 2019). "F" represents fluoride. For IQ distribution results by drinking water fluoride level provided in Xiang et al. (2003a), Trivedi et al. (2012), Sudhir et al. (2009), and Seraj et al. (2012) and rate of low IQ scores by fluoride intake provided in Wang et al. (2012), statistical significance was not evaluated.

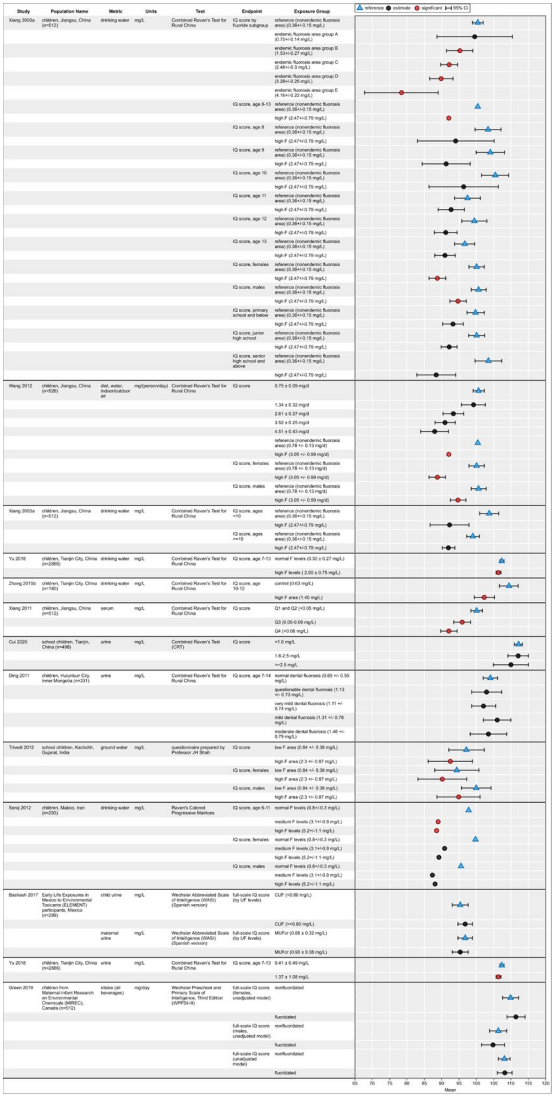


Figure A-2. Mean IQ in Children by Fluoride Exposure (Low Risk-of-bias Studies)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance. An interactive version of Figure A-2 and additional study details in HAWC here (NTP 2019) "F" represents fluoride. Three additional publications based on subsample of the larger Yu et al. (2018) cohort were identified (Zhao et al. 2019; Zhao et al. 2020; Zhou et al. 2019); 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 (NTP 2019) 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.

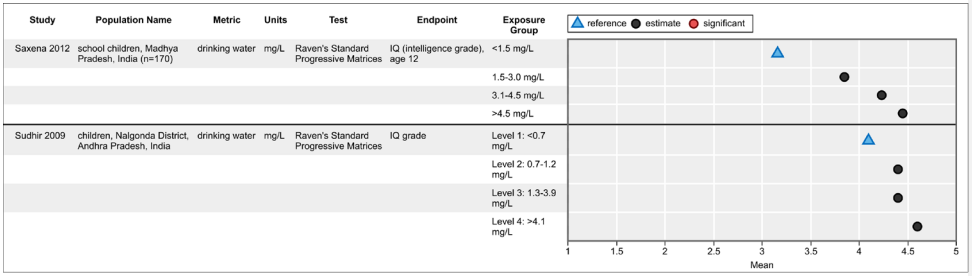


Figure A-3. Intelligence Grade in Children by Fluoride Exposure (Low Risk-of-bias Studies; Presented as Mean)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance. An interactive version of Figure A-3 and additional study details in HAWC here (NTP 2019). For Saxena et al. (2012), children's intelligence was measured using 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 et al. (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.

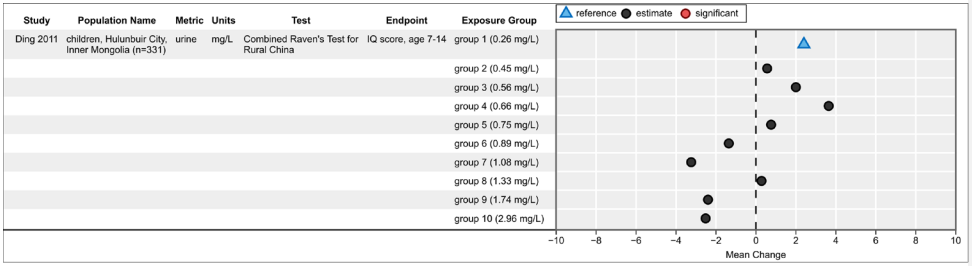


Figure A-4. Mean Change in IQ in Children by Fluoride Exposure (Low Risk-of-bias Studies)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance. An interactive version of Figure A-4 and additional study details in HAWC here (NTP 2019). For Ding et al. (2011), SDs are available and can be viewed in HAWC (NTP 2019) 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.

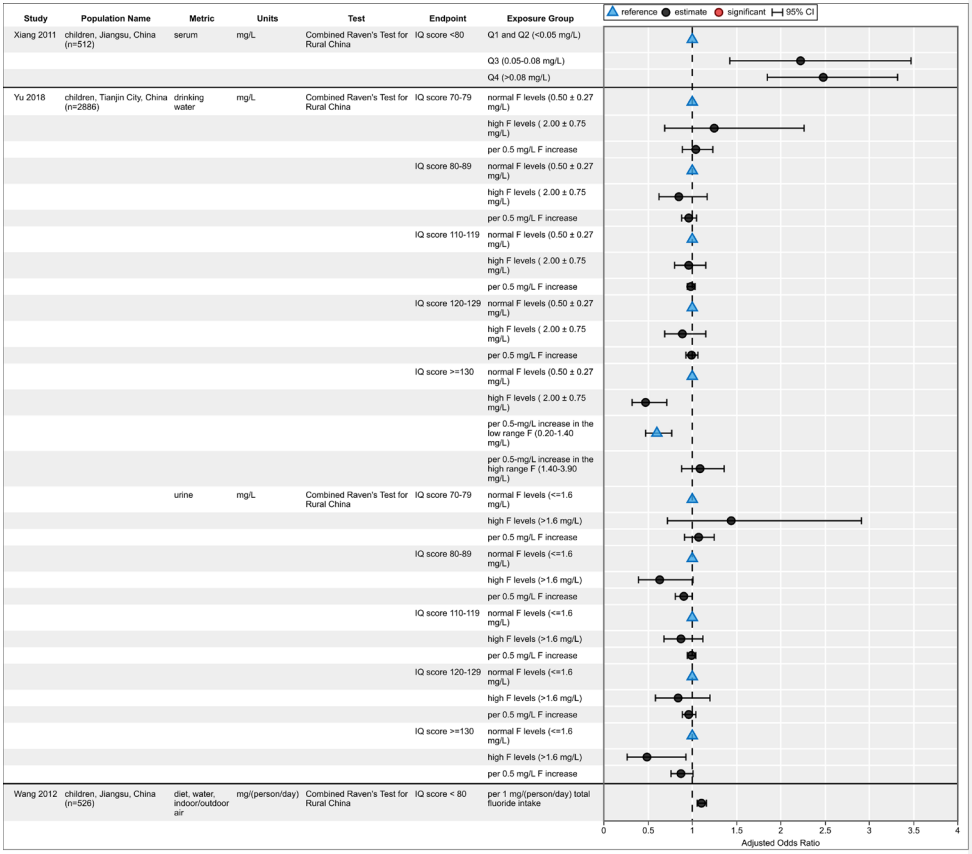
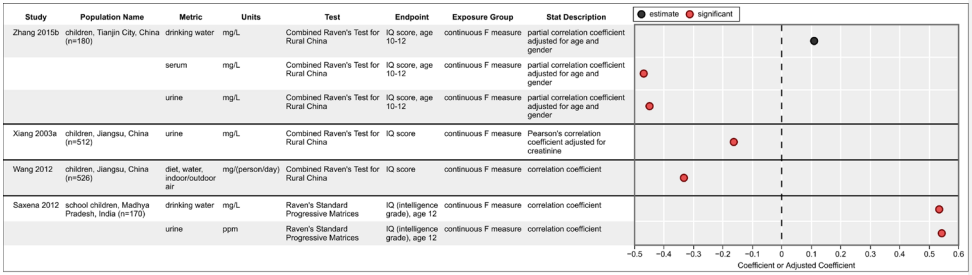


Figure A-5. Associations between Fluoride Exposure and IQ Scores in Children (Low Risk-of-bias Studies; Presented as Adjusted OR)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance. Cutoffs for the dichotomous outcome are listed in the Endpoint column.
An interactive version of Figure A-5 and additional study details in HAWC here (NTP 2019). 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.



Study	Population Name	Metric	Units	Test	Endpoint	Exposure Group	Stat Description
Zhang 2015b	children, Tianjin City, China (n=180)	drinking water	mg/L	Combined Raven's Test for Rural China	IQ score, age 10-12	continuous F measure	partial correlation coefficient adjusted for age and gender
		serum	mg/L	Combined Raven's Test for Rural China	IQ score, age 10-12	continuous F measure	partial correlation coefficient adjusted for age and gender
		urine	mg/L	Combined Raven's Test for Rural China	IQ score, age 10-12	continuous F measure	partial correlation coefficient adjusted for age and gender
Xiang 2003a	children, Jiangsu, China (n=512)	urine	mg/L	Combined Raven's Test for Rural China	IQ score	continuous F measure	Pearson's correlation coefficient adjusted for creatinine
Wang 2012	children, Jiangsu, China (n=526)	diet, water, indoor/outdoor air	mg/(person/day)	Combined Raven's Test for Rural China	IQ score	continuous F measure	correlation coefficient
Saxena 2012	school children, Madhya Pradesh, India (n=170)	drinking water	mg/L	Raven's Standard Progressive Matrices	IQ (intelligence grade), age 12	continuous F measure	correlation coefficient
		urine	ppm	Raven's Standard Progressive Matrices	IQ (intelligence grade), age 12	continuous F measure	correlation coefficient

Figure A-6. Correlations between Fluoride Exposure and IQ Score in Children (Low Risk-of-bias Studies; Presented as Correlation Coefficient)

Circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-6 and additional study details in HAWC here (NTP 2019). "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.

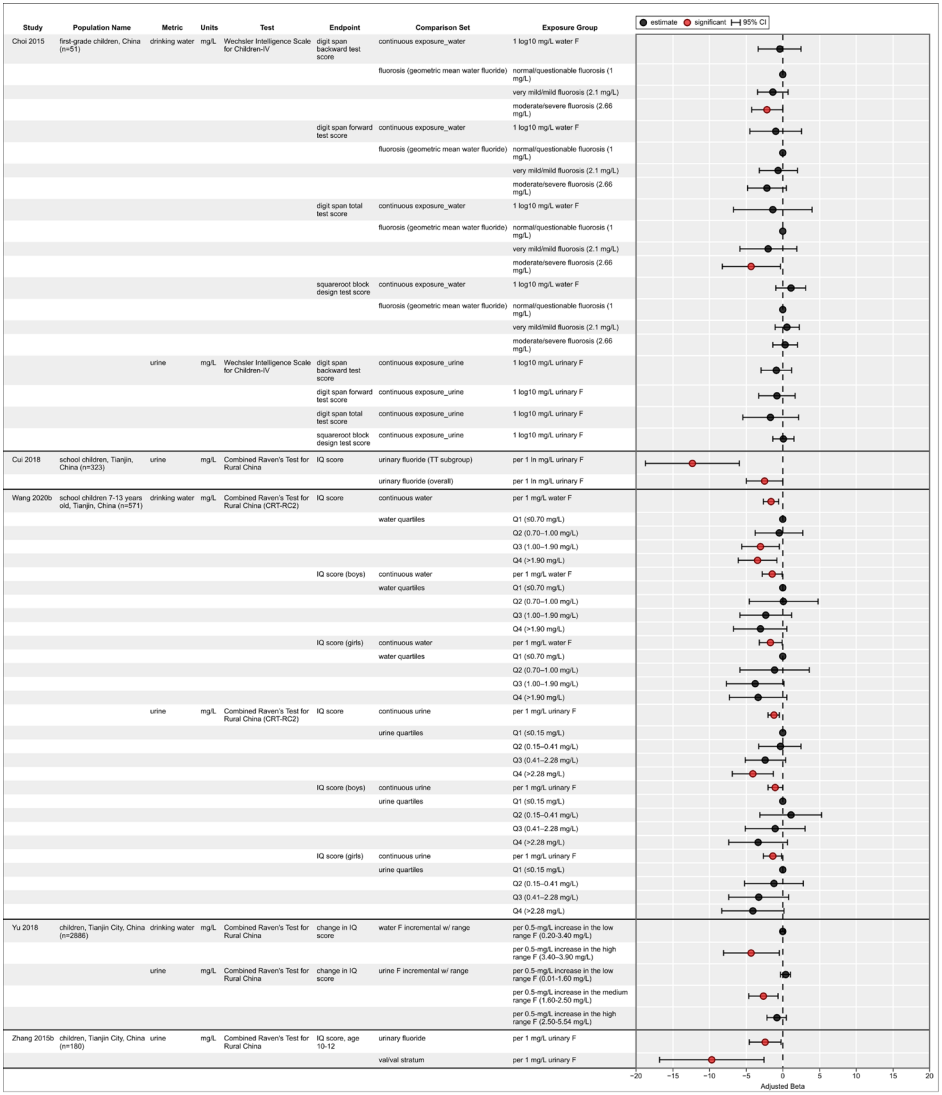


Figure A-7. Associations between Fluoride Exposure and IQ Score in Children (Low Risk-of-bias Studies; Presented as Adjusted Beta)—China

Circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-7 and additional study details in HAWC here (NTP 2019). "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 for change in IQ score were not reported.

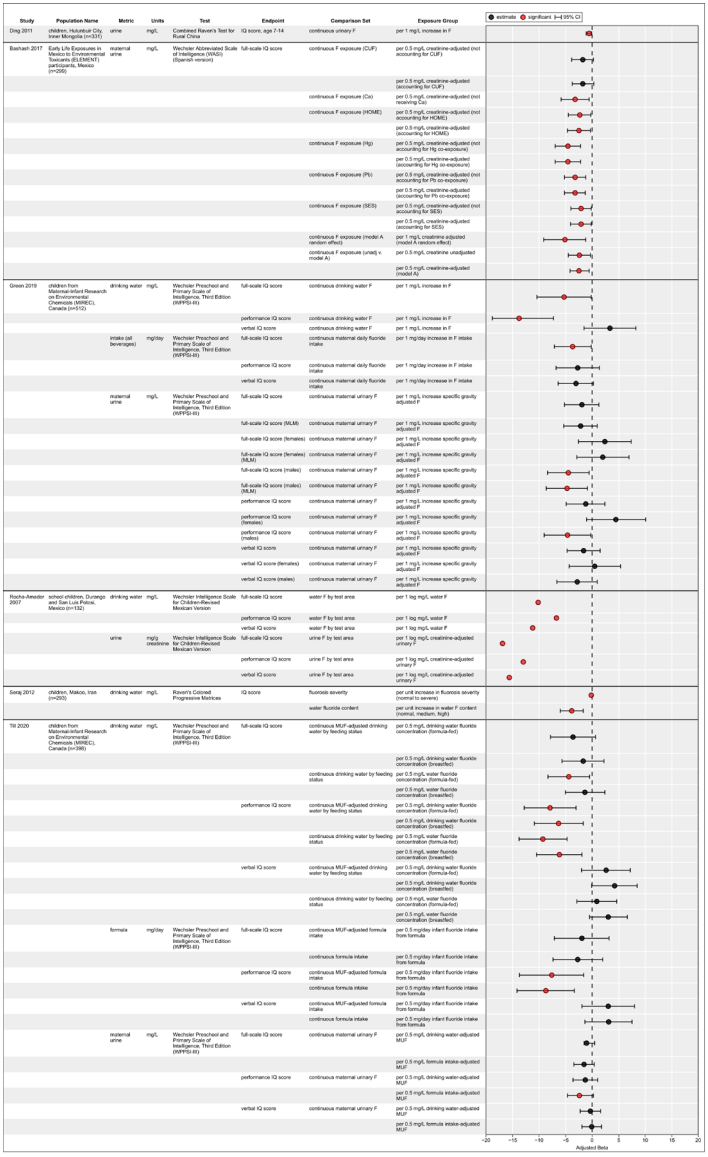


Figure A-8. Associations between Fluoride Exposure and IQ Score in Children (Low Risk-of-bias Studies; Presented as Adjusted Beta)—Areas Other Than China

Circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-8 and additional study details in HAWC here (NTP 2019). "F" represents fluoride.

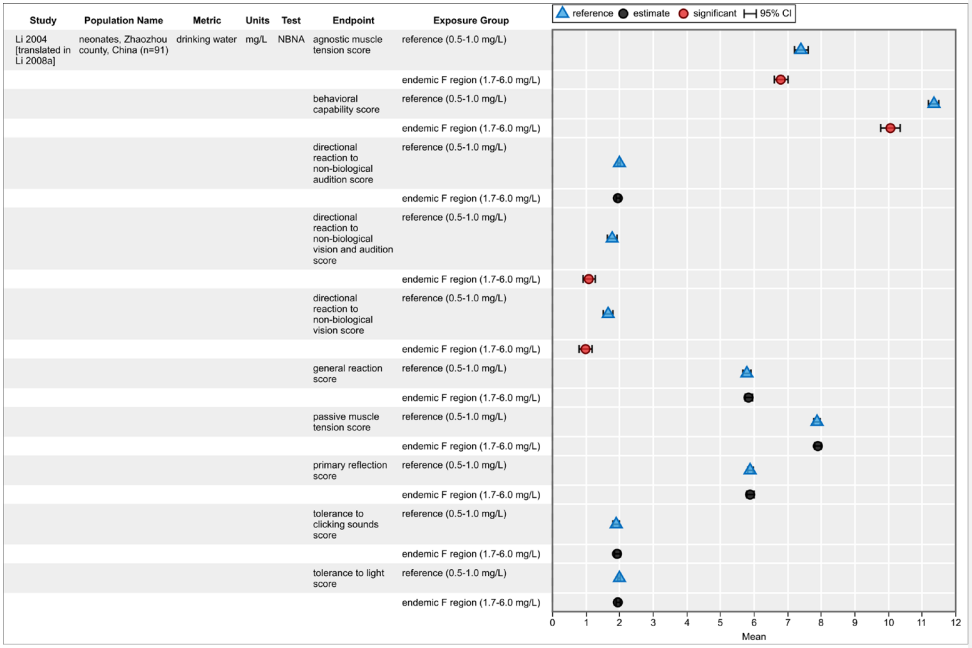


Figure A-9. Mean Motor/Sensory Scores in Children by Fluoride Exposure (Low Risk-of-bias Studies)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-9 and additional study details in HAWC here (NTP 2019). "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 (NTP 2019) by clicking the data points within the plot area. Total neonatal behavioral neurological assessment (NBNA) score was also significantly reduced in the endemic F region versus reference region (not shown).

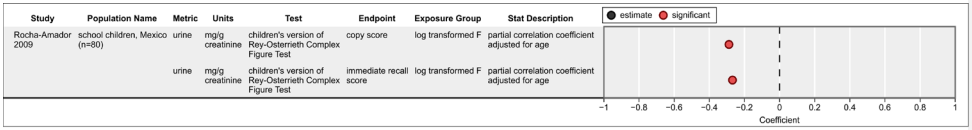


Figure A-10. Correlations between Fluoride Exposure and Other Cognitive Effects in Children (Low Risk-of-bias Studies; Presented as Correlation Coefficient)

Circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-10 and additional study details in HAWC here (NTP 2019). "F" represents fluoride.

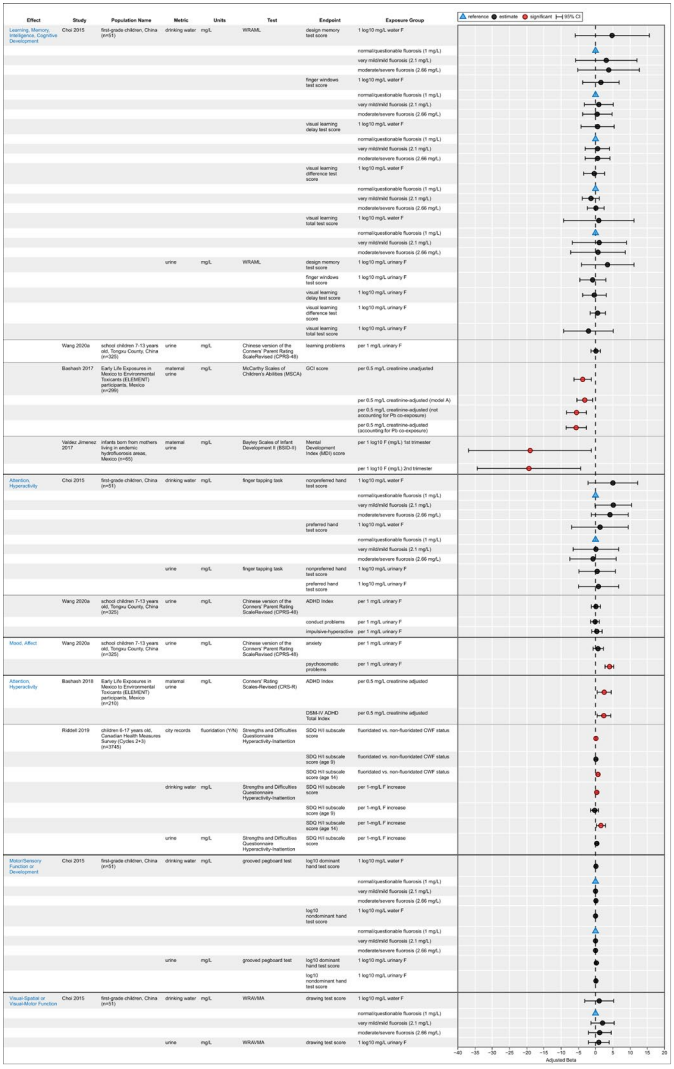


Figure A-11. Associations between Fluoride Exposure and Other Neurodevelopmental Effects in Children (Low Risk-of-bias Studies; Presented as Adjusted Beta)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance. An interactive version of Figure A-11 and additional study details in HAWC here (NTP 2019). "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.

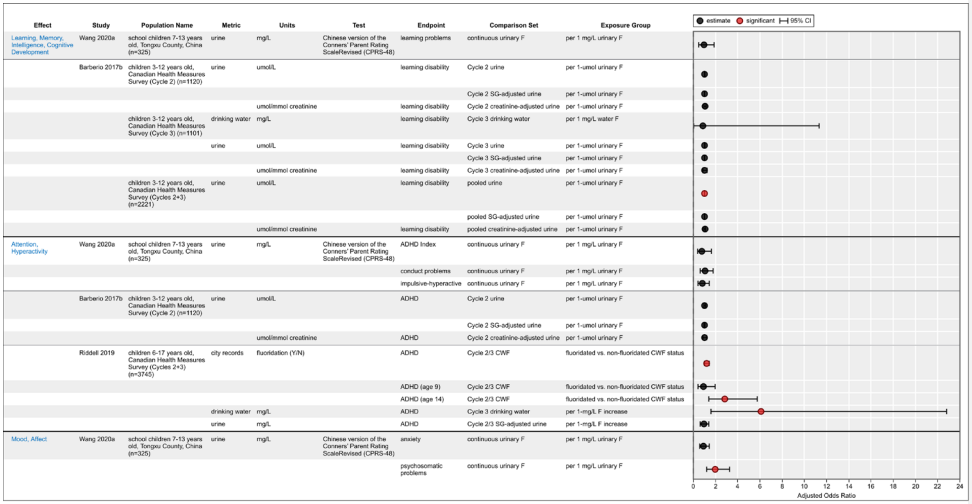


Figure A-12. Associations between Fluoride Exposure and Other Neurodevelopmental Effects in Children (Low Risk-of-bias Studies; Presented as Adjusted OR)

Circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-12 and additional study details in HAWC here (NTP 2019). "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 (NTP 2019) by clicking the OR within the plot area.

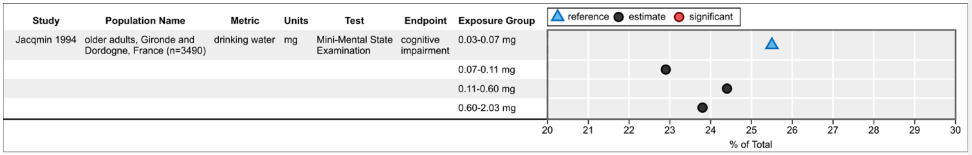


Figure A-13. Cognitive Impairment in Adults by Fluoride Exposure (Low Risk-of-bias Studies; Presented as % of Total Group)

Reference group indicated by blue triangles; circles represent response estimates with red indicating statistical significance.
An interactive version of Figure A-13 and additional study details in HAWC here (NTP 2019). 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.

Appendix B. Literature Search and Document Review Details

Table of Contents

Tables

B.1. Introduction

NTP initially published a systematic review of the experimental animal literature in 2016 that was subsequently expanded to include human epidemiological studies, mechanistic studies, and newer experimental animal literature. Table B-1 provides a timeline of key activities contributing to the 2022 NTP monograph including the multiple literature searches, draft monographs, and document review activities that have occurred since 2016.

Table B-2 is a summary of the specific search terms used for the PubMed database. In order to ensure inclusion of relevant papers, the strategy for this search was broad for the consideration of neurodevelopmental or cognitive endpoints and comprehensive for fluoride as an exposure or treatment. The specific search strategies for other databases are available in the protocol (https://ntp.niehs.nih.gov/go/785076).

Table B-1. Literature Search and Document Review Timeline

Date	Action
July 2016	Published 2016 NTP monograph of the systematic literature review on the effects of fluoride on learning and memory in animals only
June 2017	Published protocol for a new NTP monograph on systematic review on effects of fluoride on neurodevelopment and cognition from evidence in human, experimental animal, and mechanistic data
April 2019	Completed final literature search for 2019 draft NTP monograph on human, experimental animal, and mechanistic data (i.e., updated through April 2019)
May 2019	Published 2019 revised protocol for 2019 draft NTP monograph
September 2019	Sent 2019 draft NTP monograph for review by NASEM committee
February 2020	Received NASEM committee's review report of 2019 draft NTP monograph; began the following key changes in response to NASEM report: • Expanded literature search to non-English-language databases • Conducted meta-analysis on children's IQ and fluoride exposure • Revised protocol for monograph to include additional information.
May 2020	Completed final literature search for 2020 draft NTP monograph on human experimental animal and mechanistic data (i.e., updated through May 2020 and expanded to include non-English-language databases)
September 2020	Published 2020 revised protocol for 2020 draft NTP monograph
September 2020	Sent 2020 draft NTP monograph for second review by NASEM committee
February 2021	Received NASEM committee's review report of revised 2020 draft NTP monograph; made the following key changes in response to NASEM report: • Removed hazard step and hazard conclusions • Removed meta-analysis to publish separately.
December 2021	Sent 2021 draft NTP monograph on the state of the science for external peer review
April 2022	Published final 2022 NTP monograph on the state of the science

Commented [A107]: This paragraph and Table B-1 were added to provide further clarity on the progression of this multiyear assessment, in response to related comments from [redacted]Reviewers listed below; see DocI_Monograph and DocJ_Monograph, respectively, for detailed response. Note, one of the reviewers provided four separate comments on this issue.

Reviewer comment (DocI_Monograph, page 8): The rationale for date selection needs to be more clearly articulated. The specific dates are included in the appendix, perhaps they could be included in the main text for clarity in the methods.

Reviewer comment (DocJ_Monograph, page 4): Because this review has an extensive history that could be difficult for a reader to follow (i.e., the original 2016 review, and drafts from 2019, 2020, and the current draft), it would be helpful to develop a table or flowchart that documents that history. For example, you may consider noting the purpose/research question, findings, and noteworthy differences from previous/subsequent versions.
See comments below, but the literature search section, in particular, was a little difficult to follow - and having the "big picture" of the review in a table or flowchart to refer to, would better allow the reader to follow all of the searches conducted, and how the differ, yet fit together to contribute to the present document.

Reviewer comment (DocJ_Monograph, page 6): The literature search section was somewhat confusing to follow, though, given the complexity of updating reviews, etc it is understandable why multiple searches were conducted. See previous comment regarding the various iterations of this review, historically, and how a table or flowchart may help the reader understand the progression of this review, and thus, better follow the searches that were carried out.
For example, you may consider adding sub-headings within this section to distinguish which searches were run to capture which types of studies.

Reviewer comment (DocJ_Monograph, page 4): It is not clear why the "hazard assessment step" was removed from the methodology. Is it because the authors deemed the step not possible based on available evidence? Or is it because the hazard assessment step will occur separately, taking into consideration both the review and the results of meta-analysis?

Reviewer comment (DocJ_Monograph, page 10): It is not clear why the meta-analysis portion of this review is being prepared as a separate report.

Table B-2. PubMed Search Terms

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 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 C. Detailed Literature Search Results and List of Included Studies

Table of Contents

C.1. Detailed Literature Search Results

C.1.1. Literature Search Results Counts and Title and Abstract Screening

The electronic database searches retrieved 25,450 unique references in total (20,883 references during the initial search conducted in December 2016, 3,657 references during the literature search updates [including the final updated search conducted for the primary epidemiological studies on May 1, 2020], and 910 references from the supplemental Chinese database searches); 11 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, 1,036 references were moved to full-text review, and 24,425 references were excluded (11,402 by manual screening for not satisfying the PECO criteria and 13,023 based on the SWIFT-Active Screener algorithm).

C.1.2. Full-text Review

Among the 1,036 references that underwent full-text review, 489 were excluded at that stage with reasons for exclusion documented; 333 references were excluded for not satisfying the PECO criteria; and 156 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 Main 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 2) [using reporting practices outlined in ~~Moher et al. (2009)~~Page et al. (2021)]. After full-text review, 547 studies were considered relevant with primary neurodevelopmental or cognitive outcomes, secondary neurobehavioral outcomes, and/or outcomes related to thyroid function. 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:

- 167 human studies (84 primary only; 13 secondary only; 5 primary and secondary; 8 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.

Commented [A108]: Change made in response to the [redacted] Reviewer comment below; see DocB1_Monograph for detailed response.

Reviewer comment (DocB1_Monograph, page 5): The authors are using an old version of the PRISMA flow diagram - The 2020 PRISMA flow diagram can be found here: https://prisma-statement.org/prismastatement/flowdiagram.aspx

C.2. List of Included Studies

C.2.1. Studies in Humans

As described in Figure 2, 167 human studies were included; however, full data extraction was conducted only on studies with neurological outcomes or thyroid hormone data. Data extraction was completed using HAWC (NTP 2019). Data were extracted from a subset of included studies in humans (n = 124) 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 extracted only for primary neurodevelopmental or cognitive outcomes; a subset of these studies (n = 7) 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 evaluated only 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 167 human studies that were included in the review. An overview of the screening results is outlined in the study selection diagram (Figure 2) that reports numbers of included studies as well as numbers of studies excluded for each reason at the full-text review stage.

C.2.1.1. 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 F, Wang H, Yang D, Zhang Z. 2001. [Investigation on the intelligence and metabolism of iodine and fluoride in children with high iodine and fluoride]. *Chin J Control Endem Dis* 12-14.

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.

Jin T, Han T, Wei Y, Wu Y, Wang Z, Zhang, H. 2016. [Investigation on working memory level of children aged 8-12 years in coal-burning fluorosis area]. *J Environ Health* 33(5): 409-411.

Jin T, Wang Z, Wei Y, Wu Y, Han T, Zhang H. 2017. [Investigation on intelligence level of children aged 8-12 years old in coal-burning fluorosis area]. *J Environ Health* 34(3): 229-231.

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.

Liang C, Ji R, Cao J, Jiang Y, Yu B, Ma F, Wu Y, Ying B, Zhang Y, Sun S, Li Y, Emsley CL, Gao S, Hall KS, Hendrie HC. 2003. [Study on the relationship between drinking water trace elements and cognitive ability of the elderly]. *Health Res* 436-440.

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.

Luo Y, Ma R, Liu Z, Guan Z, Lou D, Zheng D. 2018. [Intelligence investigation and forensic significance of children in coal-burning fluorosis area]. *Chin J Forensic Medicine* 33(6): 590-593.

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.

Pratap SV, Singh CD, 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.

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.

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 G, Zhang M, Wang Q, Han A, Gao M, Lin P, Xiang Q. 2017. [Investigation on the relationship between serum fluoride content and IQ of children before and after reducing fluoride to water]. *Capital Public Health* 11(6): 274-277.

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 P, Cheng L. 2015. [Effect of coal-burning endemic fluorosis on children's physical development and intellectual level]. *Chin J Control Endem Dis* 30(6): 458-459.

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.

Zhao Y, Cui Y, Yu J, Zhang B, Nie J, Zhao L, Zhang Z, Liu H. 2018. [Study on the relationship between water-borne high iodine and thyroid hormone and children's intelligence level]. *J Environ Health* 35(1): 6-9.

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.

C.2.1.2. 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.

C.2.2. Studies in Non-human Animals

As described in Figure 2, 339 non-human mammal studies were included; however, full data extraction was conducted only on studies with primary neurological outcomes and/or secondary functional neurological outcomes (e.g., motor activity). Data extraction was completed using HAWC (NTP 2019). 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 and/or secondary functional neurological outcomes with relevant mechanistic data). Studies that evaluated other mechanistic endpoints, as well as studies that assessed only 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 2) that reports numbers of included studies as well as numbers of studies excluded for each reason at the full-text review stage.

C.2.2.1. 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.

C.2.2.2. 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.

C.2.3. In Vitro Experimental Studies

As described in Figure 2, 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 (NTP 2019) with the exception of in vitro studies that also reported in vivo non-human animal data that met 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).

C.2.3.1. 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.

C.2.3.2. 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 D. Risk-of-bias Figures

Figures

D.1. Studies in Humans

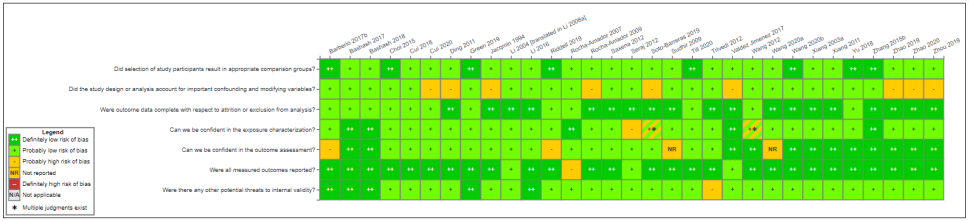


Figure D-1. Risk-of-bias Heatmap for Low Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-1 and additional study details in HAWC here (NTP 2019).

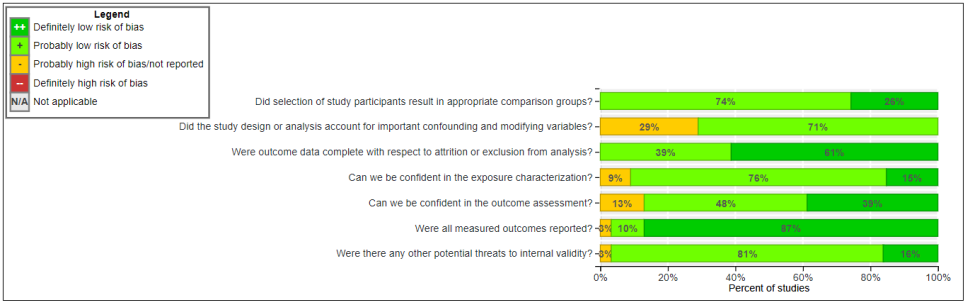


Figure D-2. Risk-of-bias Bar Chart for Low Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-2 and additional study details in HAWC here (NTP 2019).

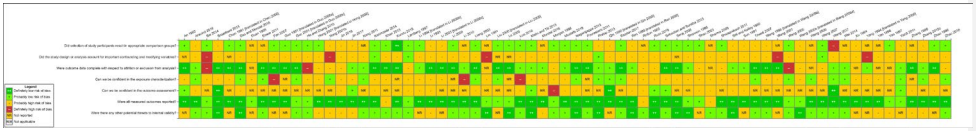


Figure D-3. Risk-of-bias Heatmap for High Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-3 and additional study details in HAWC here (NTP 2019).

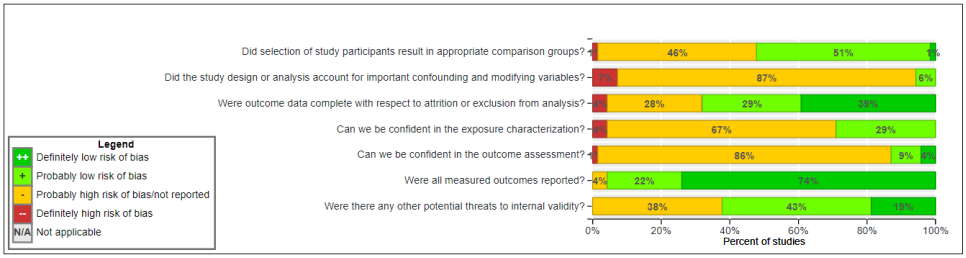


Figure D-4. Risk-of-bias Bar Chart for High Risk-of-bias Human Neurodevelopmental or Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-4 and additional study details in HAWC here (NTP 2019).

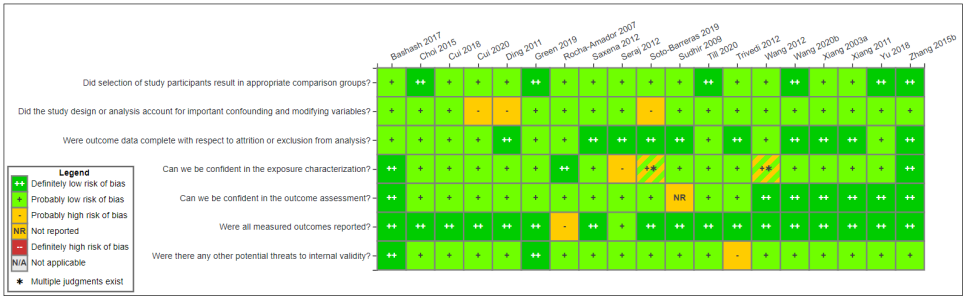


Figure D-5. Risk-of-bias Heatmap for Low Risk-of-bias Children's IQ Studies Following Fluoride Exposure

An interactive version of Figure D-5 and additional study details in HAWC here (NTP 2019).

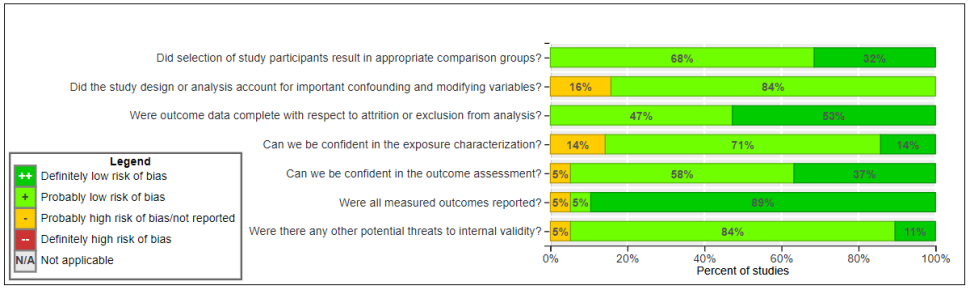


Figure D-6. Risk-of-bias Bar Chart for Low Risk-of-bias Children's IQ Studies Following Fluoride Exposure

An interactive version of Figure D-6 and additional study details in HAWC here (NTP 2019).

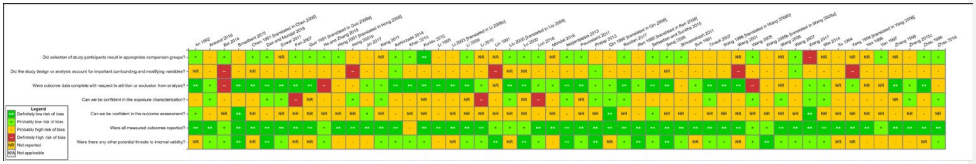


Figure D-7. Risk-of-bias Heatmap for High Risk-of-bias Children's IQ Studies Following Fluoride Exposure

An interactive version of Figure D-7 and additional study details in HAWC here (NTP 2019).

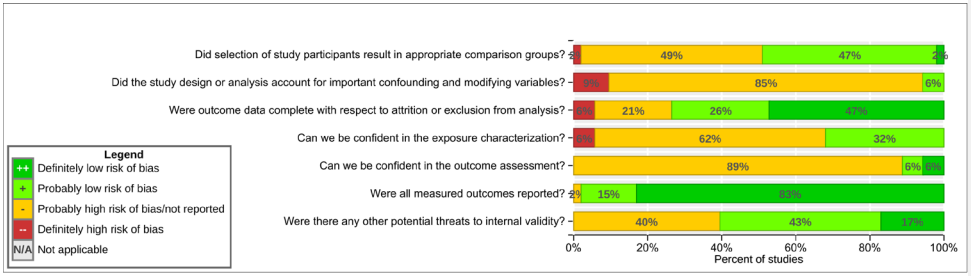


Figure D-8. Risk-of-bias Bar Chart for High Risk-of-bias Children's IQ Studies Following Fluoride Exposure

An interactive version of Figure D-8 and additional study details in HAWC here (NTP 2019).

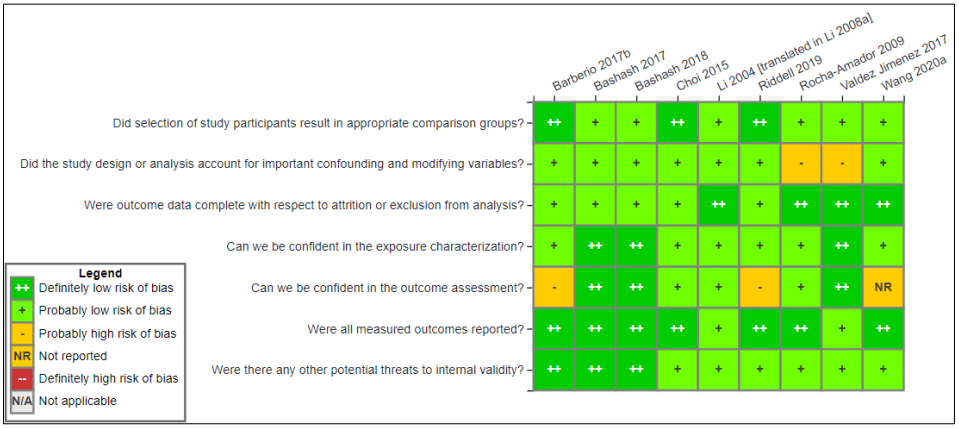


Figure D-9. Risk-of-bias Heatmap for Low Risk-of-bias Children's Other Neurodevelopmental Effect Studies Following Fluoride Exposure

An interactive version of Figure D-9 and additional study details in HAWC here (NTP 2019).

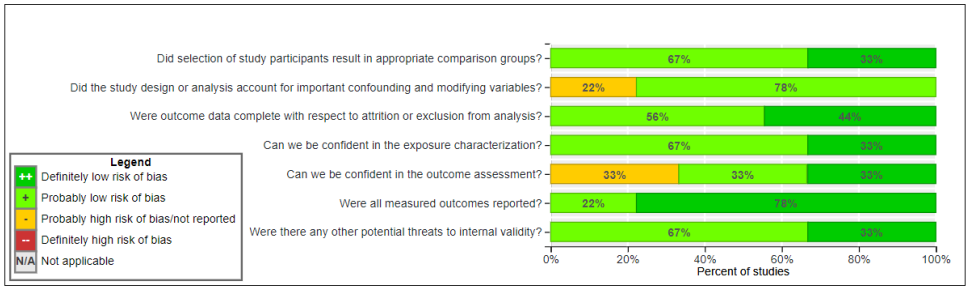


Figure D-10. Risk-of-bias Bar Chart for Low Risk-of-bias Children's Other Neurodevelopmental Effect Studies Following Fluoride Exposure

An interactive version of Figure D-10 and additional study details in HAWC here (NTP 2019).

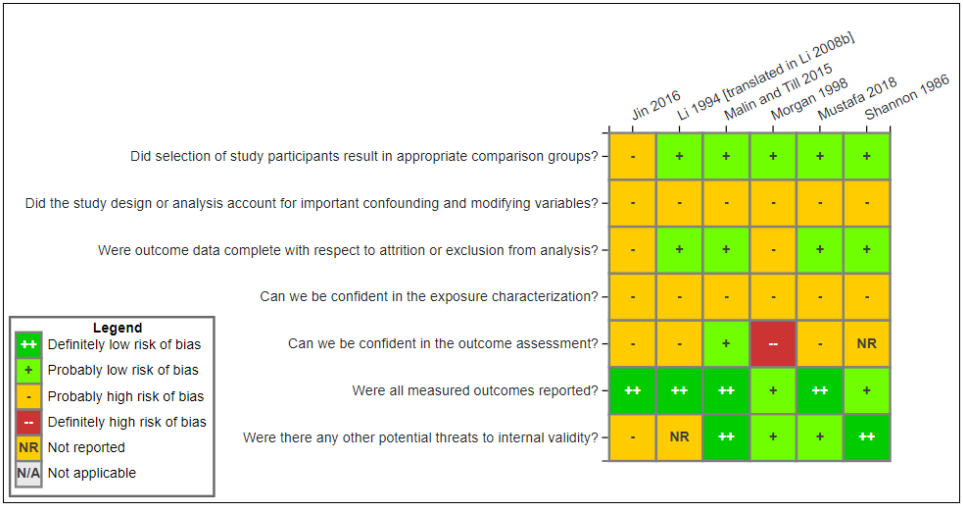


Figure D-11. Risk-of-bias Heatmap for High Risk-of-bias Children's Other Neurodevelopmental Effect Studies Following Fluoride Exposure

An interactive version of Figure D-11 and additional study details in HAWC here (NTP 2019).

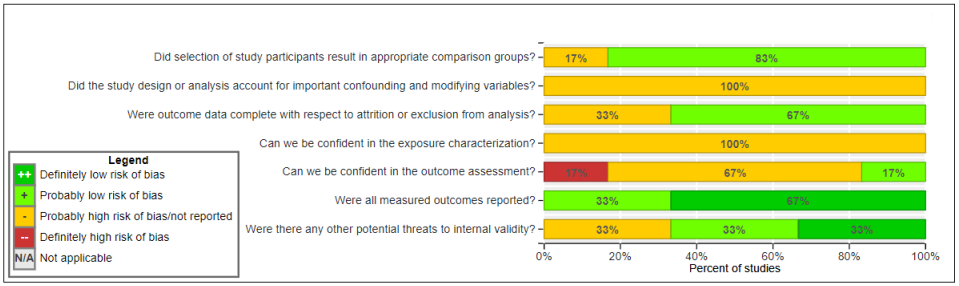


Figure D-12. Risk-of-bias Bar Chart for High Risk-of-bias Children's Other Neurodevelopmental Effect Studies Following Fluoride Exposure

An interactive version of Figure D-12 and additional study details in HAWC here (NTP 2019).

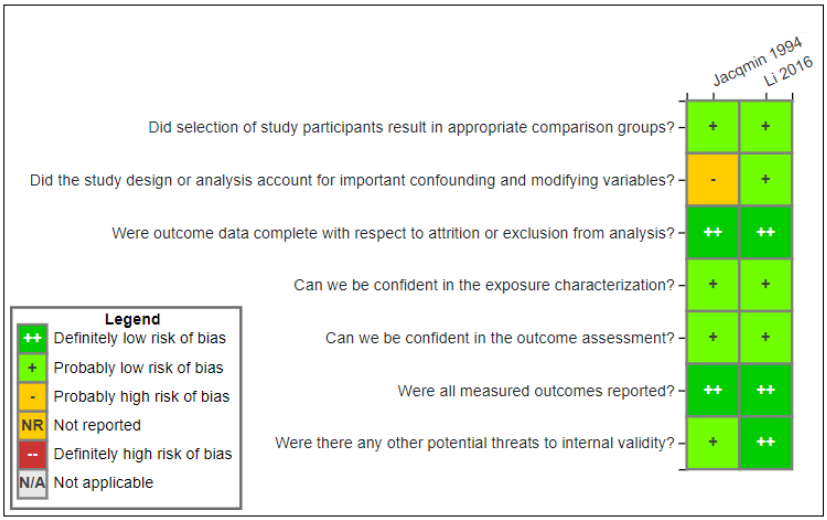


Figure D-13. Risk-of-bias Heatmap for Low Risk-of-bias Adult Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-13 and additional study details in HAWC here (NTP 2019).

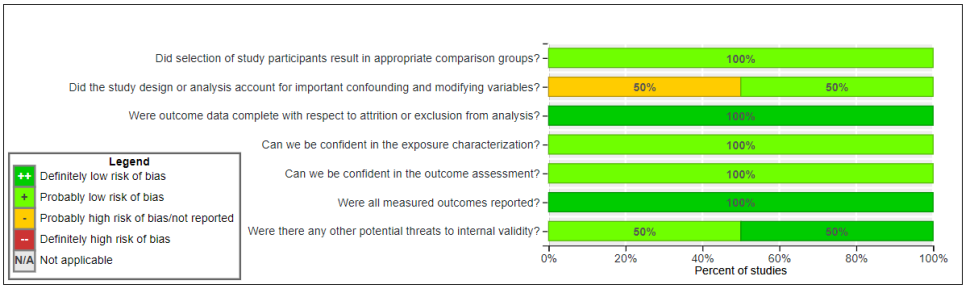


Figure D-14. Risk-of-bias Bar Chart for Low Risk-of-bias Adult Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-14 and additional study details in HAWC here (NTP 2019).

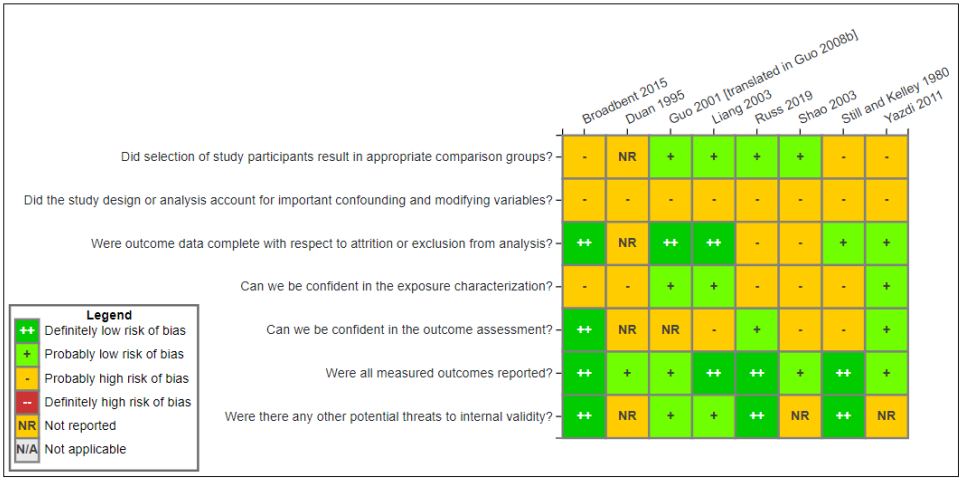


Figure D-15. Risk-of-bias Heatmap for High Risk-of-bias Adult Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-15 and additional study details in HAWC here (NTP 2019).

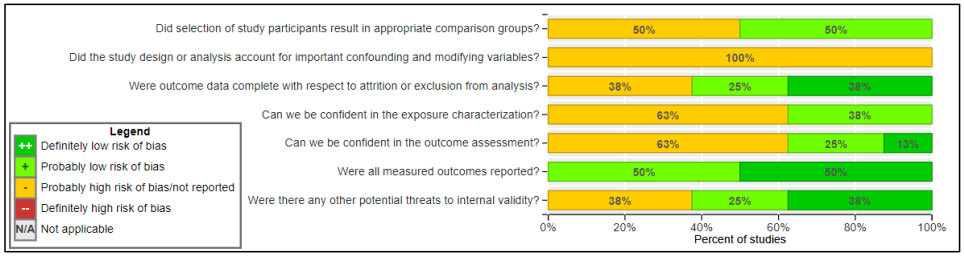


Figure D-16. Risk-of-bias Bar Chart for High Risk-of-bias Adult Cognitive Studies Following Fluoride Exposure

An interactive version of Figure D-16 and additional study details in HAWC here (NTP 2019).

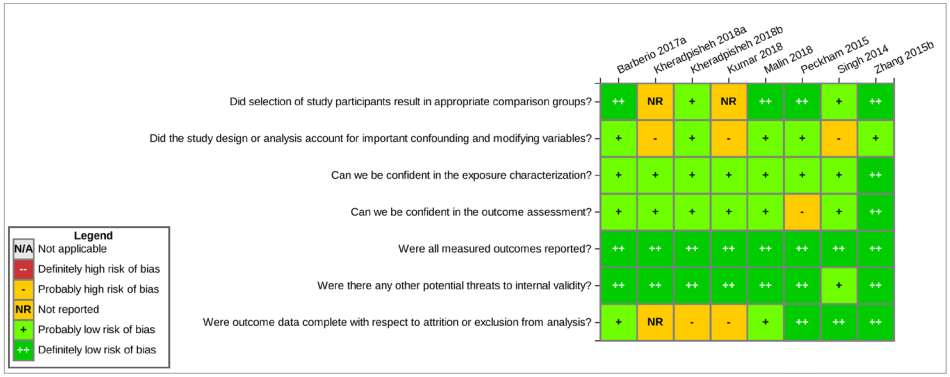


Figure D-17. Risk-of-bias Heatmap for Low Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure

An interactive version of Figure D-17 and additional study details in HAWC here (NTP 2019).

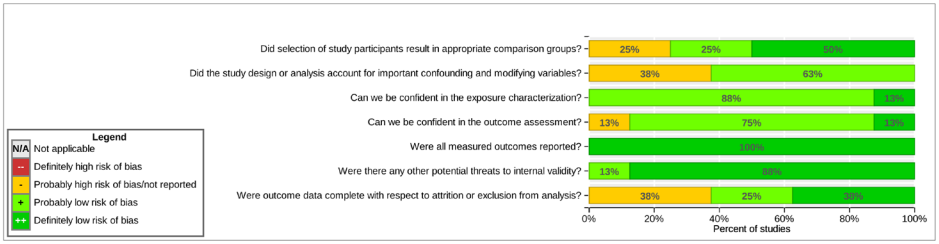


Figure D-18. Risk-of-bias Bar Chart for Low Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure

An interactive version of Figure D-18 and additional study details in HAWC here (NTP 2019).

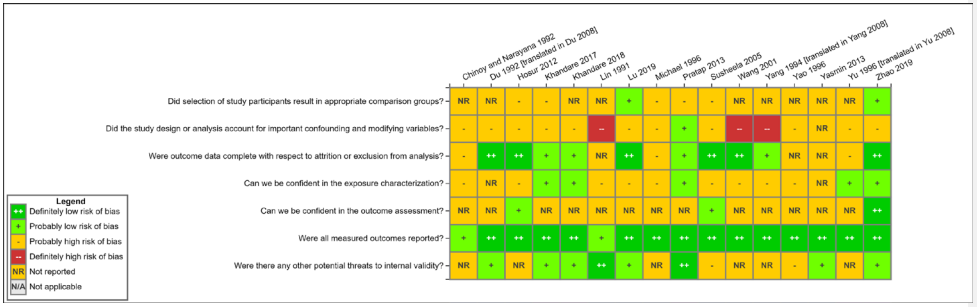


Figure D-19. Risk-of-bias Heatmap for High Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure

An interactive version of Figure D-19 and additional study details in HAWC here (NTP 2019).

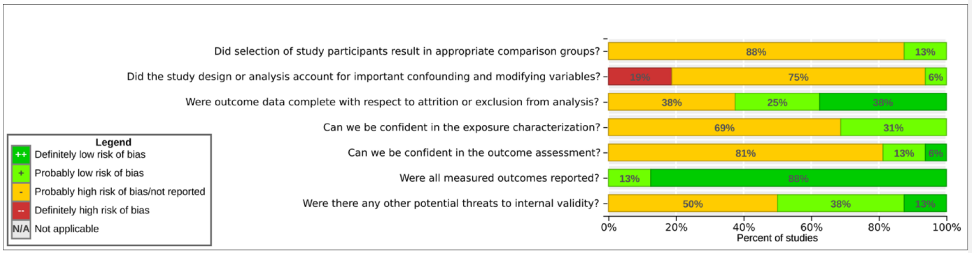


Figure D-20. Risk-of-bias Bar Chart for High Risk-of-bias Human Mechanistic Studies Following Fluoride Exposure

An interactive version of Figure D-20 and additional study details in HAWC here (NTP 2019).

D.2. Studies in Non-human Animals

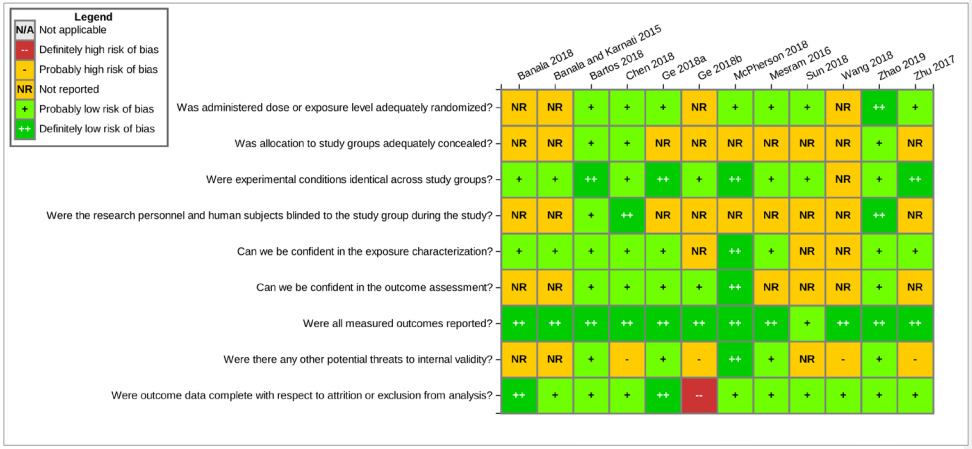


Figure D-21. Risk-of-bias Heatmap for New Developmental Animal Learning and Memory Studies Following Fluoride Exposure

An interactive version of Figure D-21 and additional study details in HAWC here (NTP 2019).

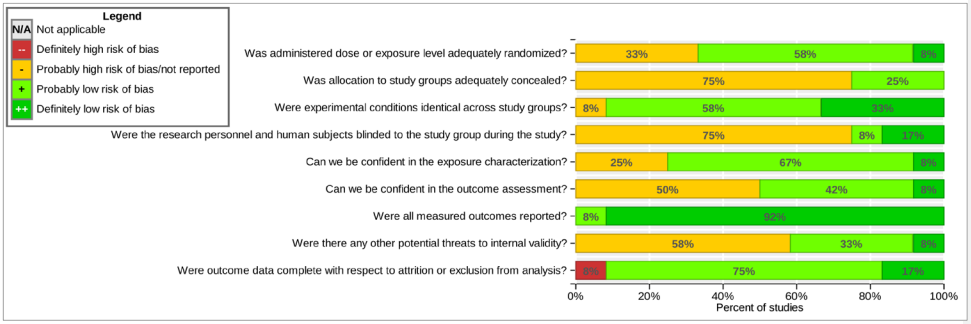


Figure D-22. Risk-of-bias Bar Chart for New Developmental Animal Learning and Memory Studies Following Fluoride Exposure

An interactive version of Figure D-22 and additional study details in HAWC here (NTP 2019).

Figure D-23. Risk-of-bias Heatmap for New Adult Animal Learning and Memory Studies Following Fluoride Exposure

An interactive version of Figure D-23 and additional study details in HAWC here (NTP 2019).

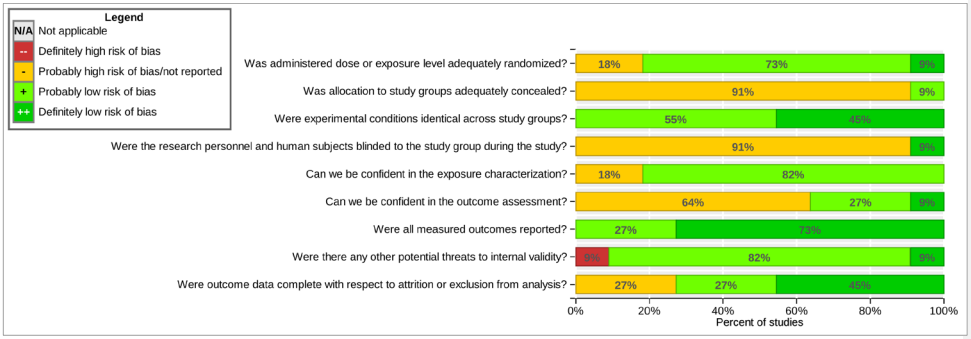


Figure D-24. Risk-of-bias Bar Chart for New Adult Animal Learning and Memory Studies Following Fluoride Exposure

An interactive version of Figure D-24 and additional study details in HAWC here (NTP 2019).

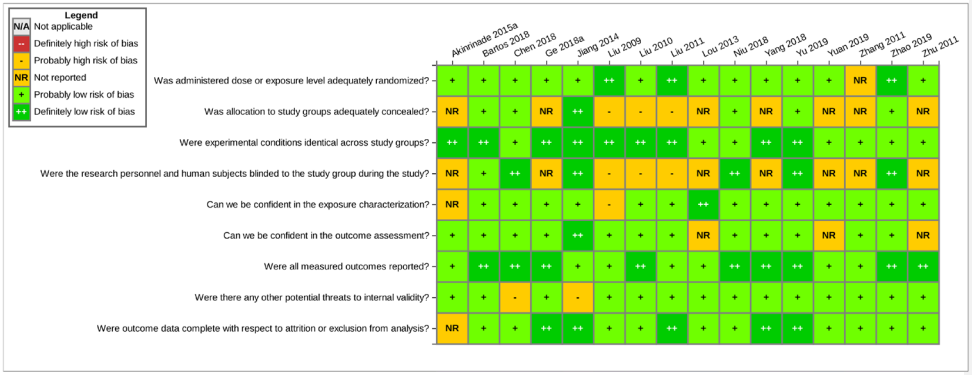


Figure D-25. Risk-of-bias Heatmap for Low Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure

An interactive version of Figure D-25 and additional study details in HAWC here (NTP 2019).

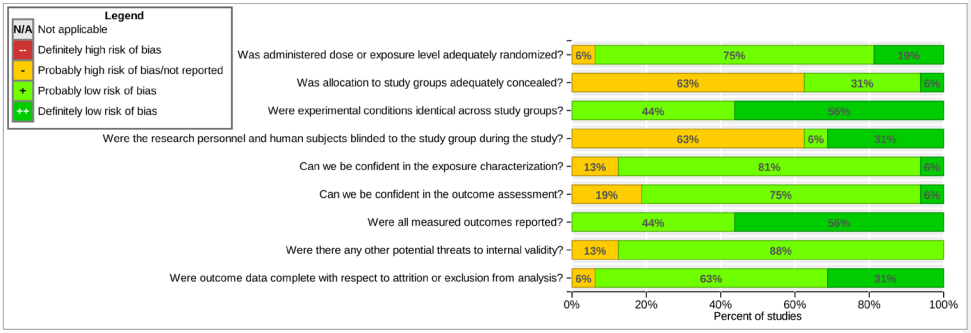


Figure D-26. Risk-of-bias Bar Chart for Low Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure

An interactive version of Figure D-26 and additional study details in HAWC here (NTP 2019).

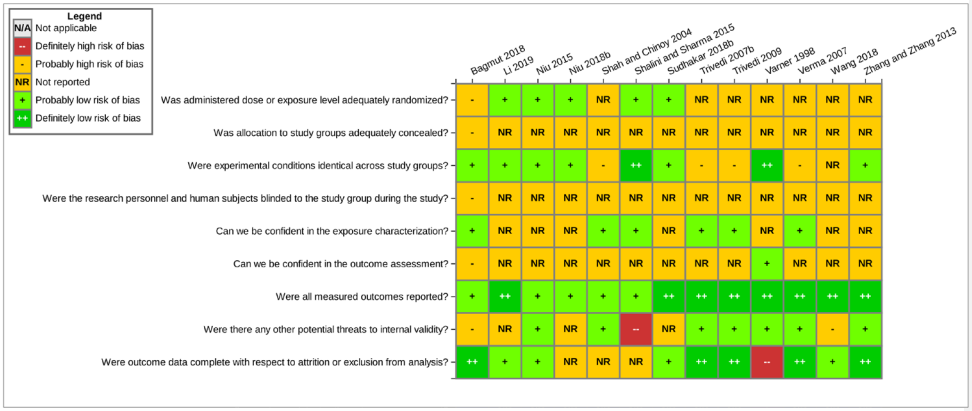


Figure D-27. Risk-of-bias Heatmap for High Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure

An interactive version of Figure D-27 and additional study details in HAWC here (NTP 2019).

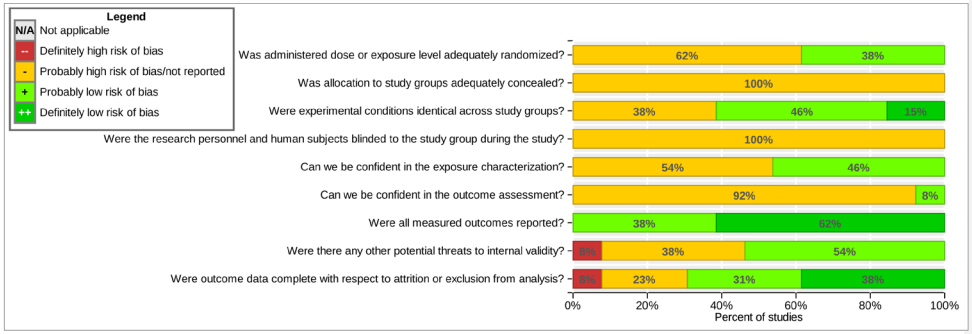


Figure D-28. Risk-of-bias Bar Chart for High Risk-of-bias Animal Biochemical Studies Following Fluoride Exposure

An interactive version of Figure D-28 and additional study details in HAWC here (NTP 2019).

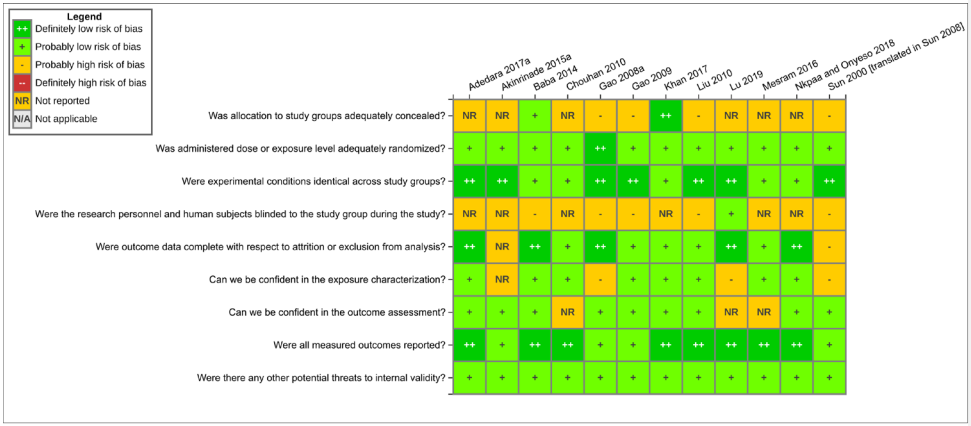


Figure D-29. Risk-of-bias Heatmap for Low Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

An interactive version of Figure D-29 and additional study details in HAWC here (NTP 2019).

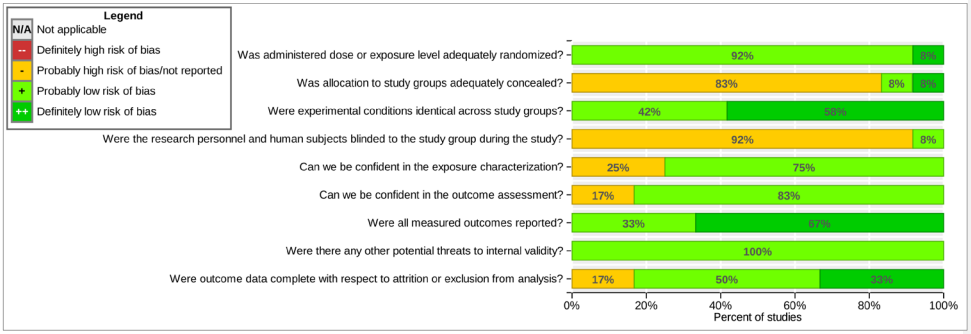


Figure D-30. Risk-of-bias Bar Chart for Low Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

An interactive version of Figure D-30 and additional study details in HAWC here (NTP 2019).

Figure D-31. Risk-of-bias Heatmap for High Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

An interactive version of Figure D-31 and additional study details in HAWC here (NTP 2019).

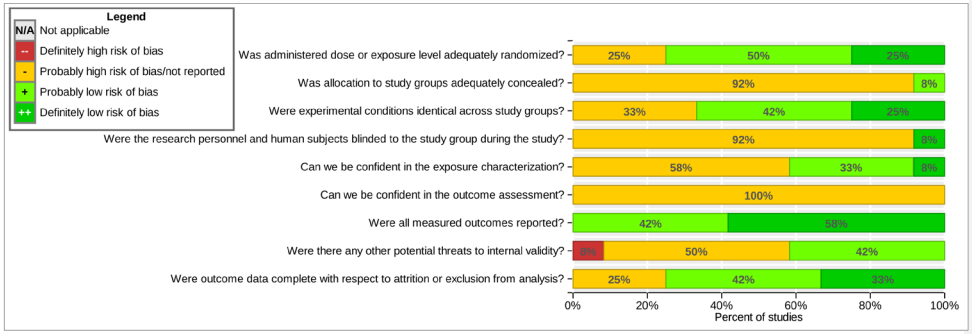


Figure D-32. Risk-of-bias Bar Chart for High Risk-of-bias Animal Neurotransmission Studies Following Fluoride Exposure

An interactive version of Figure D-32 and additional study details in HAWC here (NTP 2019).

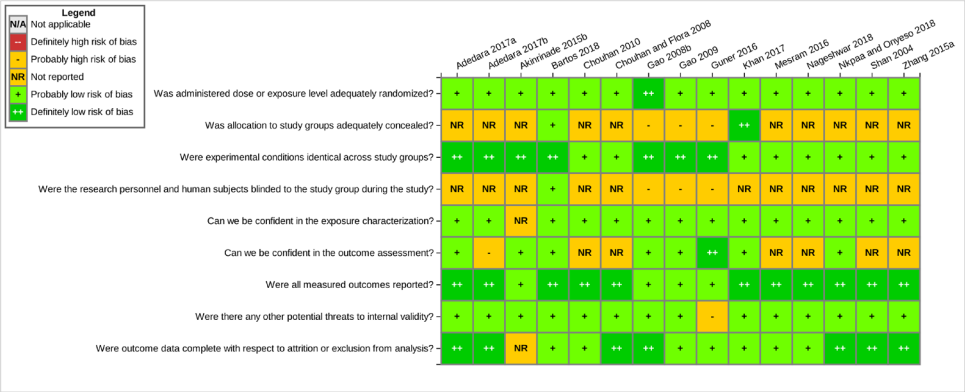


Figure D-33. Risk-of-bias Heatmap for Low Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure

An interactive version of Figure D-33 and additional study details in HAWC here (NTP 2019).

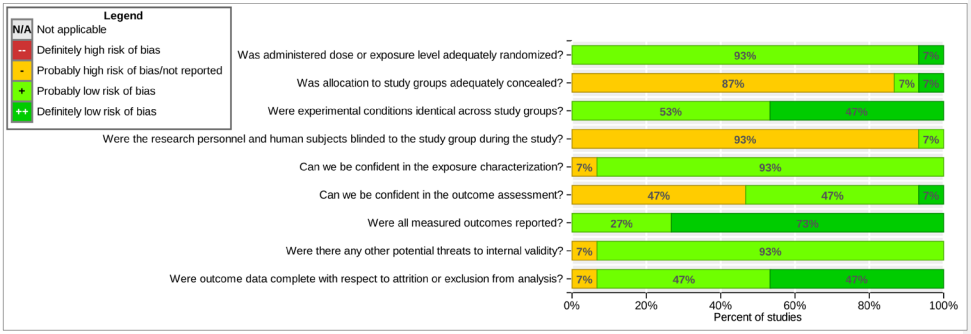


Figure D-34. Risk-of-bias Bar Chart for Low Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure

An interactive version of Figure D-34 and additional study details in HAWC here (NTP 2019).

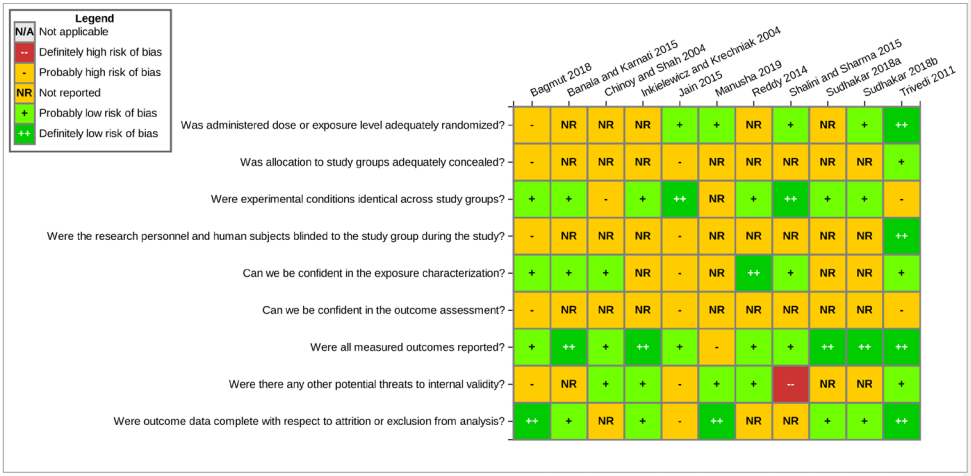


Figure D-35. Risk-of-bias Heatmap for High Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure

An interactive version of Figure D-35 and additional study details in HAWC here (NTP 2019).

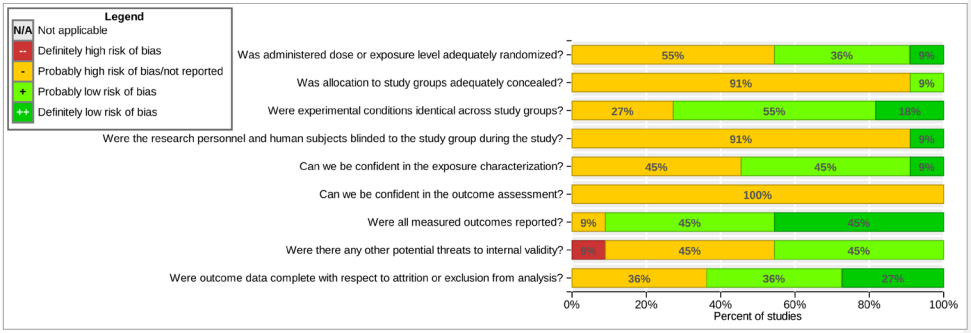


Figure D-36. Risk-of-bias Bar Chart for High Risk-of-bias Animal Oxidative Stress Studies Following Fluoride Exposure

An interactive version of Figure D-36 and additional study details in HAWC here (NTP 2019).

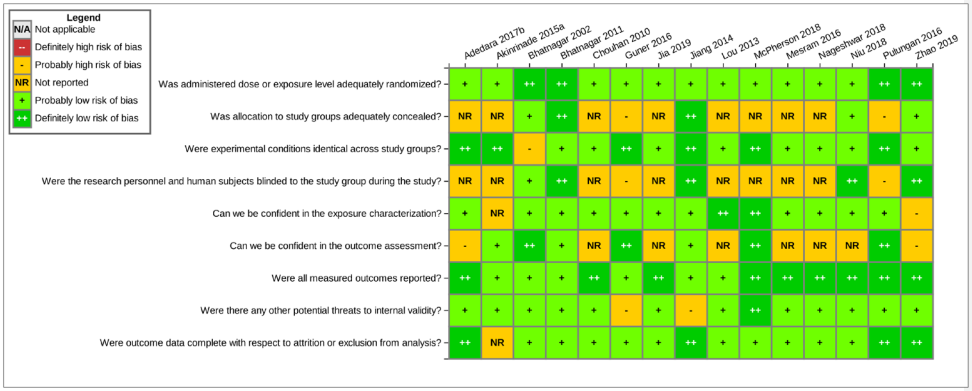


Figure D-37. Risk-of-bias Heatmap for Low Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure

An interactive version of Figure D-37 and additional study details in HAWC here (NTP 2019).

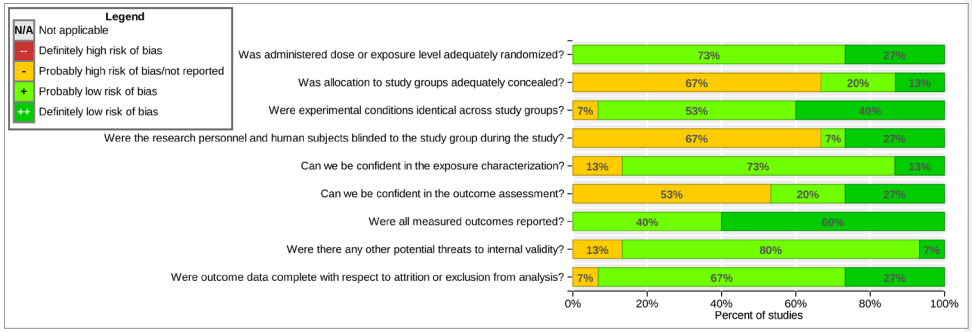


Figure D-38. Risk-of-bias Bar Chart for Low Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure

An interactive version of Figure D-38 and additional study details in HAWC here (NTP 2019).

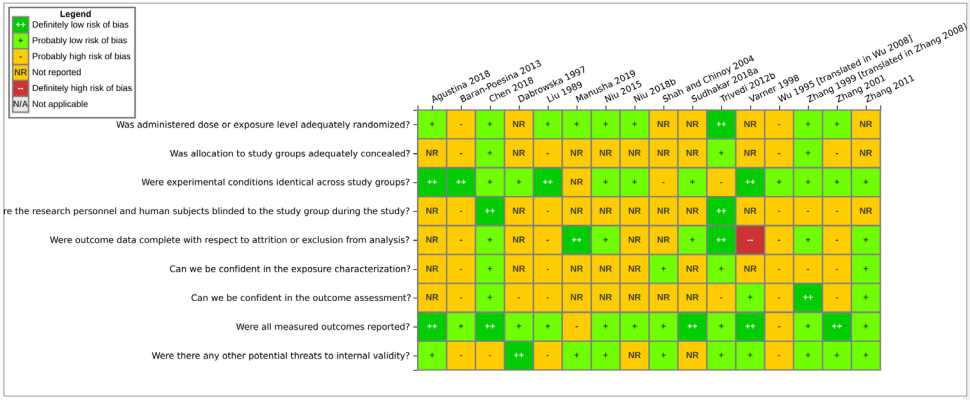


Figure D-39. Risk-of-bias Heatmap for High Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure

An interactive version of Figure D-39 and additional study details in HAWC here (NTP 2019).

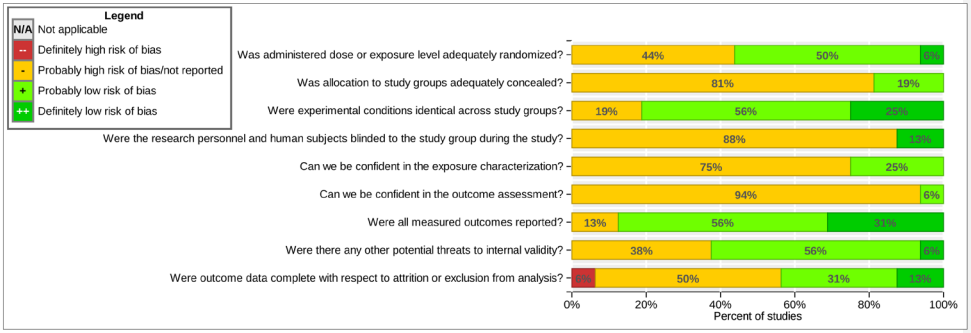


Figure D-40. Risk-of-bias Bar Chart for High Risk-of-bias Animal Histopathology Studies Following Fluoride Exposure

An interactive version of Figure D-40 and additional study details in HAWC here (NTP 2019).

Appendix E. Details for Low Risk-of-bias Studies

Table of Contents

Commented [A109]: Text in *Other potential threats* throughout Appendix E reflects revisions made to consider study-specific failures to account for sampling strategy or clustering in determining potential for bias, in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): Clustering: NASEM identified that in some population studies, participants living in the same communities were assigned the same measure of fluoride exposure without considering the effect in the data analysis. These correlations may artificially increase the statistical power.

Recommendation: Limitations should note the studies where clustering was a potential threat and specifically whether the investigators addressed this.

E.1. IQ Studies

E.1.1. Bashash et al. (2017)

E.1.1.1. 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 211 had data for the IQ analyses.
- ***Data relevant to the review:*** Adjusted and unadjusted associations between IQ scores and maternal or child's urinary fluoride concentrations.
- ***Reported association with fluoride exposure:*** Yes: Significant association between maternal urinary fluoride and IQ score (adjusted $\beta = -2.50$ per 0.5 mg/L increase; 95% CI: −4.12, −0.59). No significant associations with children's urinary fluoride.

E.1.1.2. Risk of Bias

- Author contacts:
  - Authors were contacted for additional information on whether clustering was addressed. The authors provided results from additional models with cohort as a random effect, which informed the rating decision for the following risk-of-bias domains: Other.
- 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 no information on smoking habits was considered. 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 wherein 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

Commented [A110]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

adjusted for gestational age at birth, sex, birth weight, birth order, 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. Important covariates 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 because the study authors did not discuss it as an issue, but did consider other co-exposures. Arsenic is included in the water quality control program in Mexico City and is not considered a concern in this population.

- *Potentially important study-specific covariates:* All key covariates were addressed.
  - *Direction/magnitude of effect size:* Not applicable.
- *Basis for rating:* Probably low risk of bias based on direct evidence that key covariates, 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 size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- 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. Statistical tests of bivariate associations (using Chi-square tests for categorical variables and analysis of variance [ANOVA]) were used to compare the means of the outcomes or exposure within groups based on the distribution of each covariate. Generalized additive models (GAMs) were used to estimate the adjusted association between fluoride exposure and measures of children's intelligence. Residual diagnostics were used to examine model assumptions and identify any potentially influential observations. Results are reported as adjusted effects and 95% CIs. In sensitivity analyses, regression models accounted for clustering at the cohort level by using cohort as a fixed effect in the models. Although using cohort as a random effect would be more appropriate, using individual-level exposure data and accounting for numerous important covariates in the models likely captured the cohort effect. Additional models with cohort as a random effect were also subsequently made available via personal communication with the study authors and showed similar results to the main model.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Definitely low risk if bias is based on direct evidence that the statistical analyses were appropriate and there were no other potential threats of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include

individual exposure measurements, blinding of outcome assessor to participants' fluoride exposure, and the prospective cohort study design.

E.1.2. Choi et al. (2015)

E.1.2.1. 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 IQ (digit span for auditory span and working memory and block design for visual organization and reasoning components of WISC-IV only) 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 (adjusted β = −2.13; 95% CI: −4.24, −0.02) digit span scores. Linear associations between total digit span and log-transformed 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) were observed but not significant. Other outcomes not significantly associated with fluoride exposure.

E.1.2.2. 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 whether the 51 children represented all the first-grade children from this area or whether 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. Important covariates 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 have been used as a covariate of neurodevelopmental performance. Important covariates 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 have been a concern.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is direct evidence that the key covariates were considered and indirect evidence that co-exposure to arsenic was 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 totals only 43, but the text indicates 8 children did not have a Dean Index because permanent 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 analytic methods were not reported, but it is likely that standard methods were used because the

analyses 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 size:* 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) includes 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 the tests were validated on the study population, the study authors indicated that the tests were culture-

independent (+ for methods). Blinding of the outcome assessors to participants' fluoride exposure, or steps to minimize potential bias were 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 using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.

- 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 are appropriate. Multiple regression models evaluate the associations between exposure indicators and test scores after adjusting for covariates. Specific regression models are not described or refenced, just stated to be "standard regression analysis with confounder adjustment." The distributions of fluoride concentrations in urine and water are skewed and log10-transformed to approximate a Gaussian distribution (test not specified). Results are reported as adjusted effects and 95% CIs. There is no evidence that residual diagnostics were used to examine model assumptions; however, the impact on the effect estimates is expected to be minimal.
    - *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:* Probably low risk if bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in the confounding, exposure, and outcome domains. Study strengths include individual fluoride measurements with blinding at outcome assessment likely. All key covariates and many other important covariates were considered in the study design or analysis.

E.1.3. Cui et al. (2018)

E.1.3.1. 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 association between IQ score and log-transformed urinary fluoride (adjusted β = −2.47; 95% CI: −4.93, −0.01).

E.1.3.2. Risk of Bias

- Author contacts:
  - Authors were contacted in June 2019 to obtain additional information for risk-of-bias evaluation. Additional information provided by the authors informed the rating decision for the following risk-of-bias domains: Detection (outcome assessment).
- 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. Covariates that were assessed include: sex, age, child's ethnicity, 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,

Commented [A111]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

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 were no other environmentally toxic substances that might have affected intelligence, such as high arsenic or iodine deficiency according to the Tianjin Centers for Disease Prevention and Control.

- *Potentially important study-specific covariates:* All key covariates were considered in this study.
  - *Direction/magnitude of effect size:* Not applicable.
- *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key covariates were considered, methods for collecting the information were valid and reliable, and co-exposure to arsenic was 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 whether these children were from the same schools or whether they were evenly distributed throughout the study area. It is also unclear whether 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 during 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 size:* Not accounting for dilution could cause 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 was unlikely that the outcome assessor would have knowledge of the child's urine fluoride levels, there was potential that they would know whether 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 status of participants. However, there was 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes 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 appropriate. Multiple linear regression models were applied to evaluate the relationship between urine fluoride levels and IQ scores, accounting for numerous important covariates. The urinary fluoride levels were log-transformed due to a skewed distribution. Residual diagnostics were used to examine model assumptions. Model robustness was tested through bootstrap, sensitivity analysis after excluding potential outliers, and cross-validation techniques. Results are reported as adjusted effects and 95% CIs. The analysis did not account for clustering at the school level or at the grade level (students were from four schools in grades selected via a clustered sampling method). There is no evidence that the sampling strategy was otherwise accounted for via sampling weights. The impact of these factors on the effect estimates is expected to be minimal given the use of individual-level data and adjustment for several important covariates.
    - *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 of risk of bias identified.

- ***Basis for classification as low 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 covariates were considered in the study design or analysis.

E.1.4. Cui et al. (2020)

E.1.4.1. 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) study
- ***Sample size:*** 498 school children
- ***Data relevant to the review:*** IQ scores by urine fluoride levels.
- ***Reported association with fluoride exposure:*** Yes: A 2-point decrease in IQ was observed in the highest urinary fluoride group compared to the lowest urinary fluoride group (i.e., 110.00 in ≥2.5-mg/L group versus 112.16 in <1.6-mg/L group); however, the results did not achieve statistical significance based on a one-way ANOVA comparing the three different urinary fluoride categories only.

E.1.4.2. 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. Additional information provided by the authors regarding Cui et al. (2018) informed the rating decision for the following risk-of-bias domains: Detection (outcome assessment). ~~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

Commented [A112]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

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 covariates, 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 whether these varied by fluoride level. However, Cui et al. (2018) indicate that stress and anger were not significantly associated with fluoride, and it was assumed that results would be similar for this study even though more children were included.
  - *Potentially important study-specific covariates:* Age (children 7–12 years old)
    - *Direction/magnitude of effect size:* Age is a key covariate for IQ, even in the narrow age range evaluated in this study. The direction of the association 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 likely bias the association away from the null.
  - *Basis for rating:* Probably high risk of bias because there is indirect evidence that age was not addressed as a key covariate and it may be related to both IQ and exposure.
- Attrition:
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* Of the 620 children recruited, 122 (20%) were excluded due to incomplete information or inadequate blood sample. No information was provided to indicate whether 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 size:* Not accounting for dilution could cause 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 had knowledge of the child's urinary fluoride. Subjects appear to have been recruited based on iodine levels; therefore, it is unlikely that there would have been 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes 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:* One-way ANOVA was used to make comparisons between mean IQ by urinary fluoride levels. Consideration of heterogeneity of variances was not reported. There is no adjustment for covariates or for clustering of children at the school level. There is no evidence that the sampling strategy was otherwise accounted for (i.e., via sampling weights). The impact of these factors on the effect estimates is expected to be minimal given the use of individual-level data. The primary focus of the study was to evaluate associations between IQ and thyroid hormone or dopamine levels (not between IQ and fluoride levels). It should also be noted that more advanced analyses used for thyroid hormone- and dopamine-IQ associations

still lacked adjustment for school and accounting for clustering of children from the same school.

- *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 of risk of bias identified.

- ***Basis for classification as low 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 lack of addressing age as a key covariate.

E.1.5. Ding et al. (2011)

E.1.5.1. 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 1-mg/L increase in urinary fluoride was associated with a decrease in IQ score of 0.59 points; 95% CI: −1.09, −0.08).

E.1.5.2. 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. Details were not provided, except for a reference supporting the statement. In addition, iodine deficiency was noted as not being issue in any of the four areas. Age was the only key covariate adjusted for in the regression model. Although dental fluorosis severity by % female was reported, not enough data were provided to determine whether sex should have been considered in the regression model. 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 covariates:* Sex
    - *Direction/magnitude of effect size:* There is not enough information to determine whether there was an effect from sex. There were some differences in dental fluorosis level by sex, but it is unclear how this might impact the results or whether the distribution of sex differed by age.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that there were differences in sex 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 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 size:* Spot urine samples that did not account for dilution could have exposure misclassification. The misclassification is likely non-differential, and the 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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 and multiple linear regression), but consideration of homogeneity of variance was not reported. The NASEM committee's review (NASEM 2021) pointed out a potential concern regarding the lack of accounting for clustering at the school level because children were selected from four elementary schools. However, as outlined 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 to the extent possible and that the four elementary schools appeared to be very similar in teaching quality. There is no evidence that the sampling strategy was otherwise accounted for (i.e., via sampling weights). The impact of these factors on the effect estimates is expected to be minimal given the use of individual-level data and adjustment for age as a key covariate.
    - *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 of risk of bias.
- ***Basis for classification as low 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 lack of consideration of sex as a key covariate.

E.1.6. Green et al. (2019)

E.1.6.1. 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 sexes together and separately, with maternal urinary fluoride across all three trimesters or with estimated maternal fluoride intake.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower full-scale IQ per 1-mg/L increase in maternal urinary fluoride in boys (adjusted β = −4.49) but not girls (adjusted β = 2.40) and not in both sexes combined (adjusted β = −1.95); significantly lower full-scale IQ per 1-mg increase in maternal intake in both sexes combined (adjusted β = −3.66 [no sex interaction]); significantly lower full-scale IQ per 1-mg/L increase in drinking water fluoride in both sexes combined (adjusted β = −5.29 [no sex interaction]).

E.1.6.2. Risk of Bias

- Author contacts:
  - Authors were contacted in June 2019 for additional information for the risk-of-bias evaluation. Additional information provided by the authors informed the rating decision for the following risk-of-bias domains: Other.
- Population selection:
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* Pregnant women were recruited from the same population during the same time frame 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, sex, age at

Commented [A113]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

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 covariates:* All key covariates were addressed.
  - *Direction/magnitude of effect size:* 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 covariates, 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 size:* 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 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 of whether the city had fluoridated water.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the outcome was assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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:* Multivariate linear regression analyses were used to evaluate the associations between maternal urinary fluoride and fluoride intake and children's IQ scores. Regression diagnostics were used to test assumptions for linearity, normality, and homogeneity. There were no potential influential observations (based on Cook's distance). Sensitivity analyses showed that the effects of maternal urinary fluoride (MUF), fluoride intake, and water fluoride were robust to the exclusion of two very low IQ scores in males (<70). City was accounted for as a covariate in the regression models published. Additional models with city as a random effect were also subsequently made publicly available and showed similar results to the main model.
    - *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 of risk of bias identified.
- ***Basis for classification as low 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 consideration of key covariates.

E.1.7. Rocha-Amador et al. (2007)

E.1.7.1. 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 log-transformed fluoride and IQ scores (full-scale IQ adjusted βs of −10.2 [water] and −16.9 [urine]; CIs not reported); arsenic also present, but the effect from arsenic was smaller (full-scale IQ adjusted βs of −6.15 [water] and −5.72 [urine]; CIs not reported).

E.1.7.2. Risk of Bias

- Author contacts:
  - Authors were not contacted for additional information to inform the risk-of-bias evaluation 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, sex, or time of residence were observed between participants and non-participants. Time frame 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 considered 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 age, sex, SES, transferrin saturation, weight, height, blood lead levels, and mother's education. Arsenic

Commented [A114]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response.

Reviewer comment (DocA1_Monograph, page 3): The NTP re5ponse to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

levels were highly correlated with fluoride levels; however, arsenic and fluoride were evaluated alone, and arsenic was found to have less of an effect on IQ than fluoride. This provides evidence that arsenic had been addressed as a co-exposure and cannot explain the association between fluoride exposure and decreased IQ. Smoking was not addressed and methods for measuring many of the covariates were not reported.

- *Potentially important study-specific covariates:* Arsenic
  - *Direction/magnitude of effect size:* The presence of arsenic in this study, which also demonstrated an association, would likely bias the association away from the null. Although arsenic may contribute to some of the magnitude of the observed effect of fluoride (the exact impact of arsenic on the magnitude cannot be assessed), the presence of arsenic does not fully explain the observed association between fluoride exposure and IQ. The presence of arsenic may affect the magnitude of the association between fluoride and IQ, but it has no impact on the direction of the association.
- *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 covariates 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, sex, or time of residence between responders and non-responders. However, no data were provided to support this, and no breakdown of responders/non-responders by region was 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 and 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 use the tap water only 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 size:* 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 was 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Probably high risk of bias (−)
  - *Summary:* It was reported that an interaction between fluoride and arsenic was measured, but it was noted only 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:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used were appropriate for the study. Multivariate linear analyses were used to evaluate the associations between fluoride in water and urine and children's IQ scores. Exposures were natural log-transformed, but the rationale was not provided. Regression diagnostics were not used to test model assumptions for linearity, normality, and homogeneity. The analyses did not account for clustering at the community level. The three selected communities were similar in population and general demographic characteristics. Although the analysis used individual-level exposures rather than area-level exposures, if the exposure levels within a certain area were highly correlated (which might be expected), then the results might still be biased. However, the overall impact on the effect estimates is expected to be minimal given the use of individual-level data and adjustment for multiple important covariates.

    - *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 of risk of bias identified.
- ***Basis for classification as low 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 blinding of outcome assessors to participants' fluoride exposure, but it is limited by the cross-sectional study design and the inability to completely rule out the influence of arsenic in the results.

E.1.8. Saxena et al. (2012)

E.1.8.1. 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; higher IQ grades are associated with lower intelligence) by water fluoride quartiles, continuous water fluoride, or continuous urinary fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant correlations between IQ score and water (r = 0.534) and urinary (r = 0.542) fluoride levels. Significant increase in mean IQ grade (i.e., increase in proportion of children with intellectual impairment) with increasing urinary fluoride in adjusted analyses.

E.1.8.2. Risk of Bias

- Author contacts:
  - Authors were contacted in August ~~of~~ 2017 to obtain additional information for risk-of-bias evaluation. Additional information provided by the authors informed the rating decision for the following risk-of-bias domains: Detection (outcome assessment).
- 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 whether the 170 subjects were selected with 100% participation or whether the 170 subjects were all who were asked to participate, but it appears that all subjects participated.

Commented [A115]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

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 covariates, 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 the 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 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. 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 increased with increasing fluoride even though there was no significant difference by group.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* 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 covariates, 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. Each subject was also asked to collect a sample of his/her first morning urine. The

fluoride content in the urine was determined using a fluoride ion-selective electrode. QA/QC and LOD were not reported, and urinary dilution was not assessed. Although only current levels were measured, children who had changed their water source since birth were excluded.

  - *Direction/magnitude of effect size:* 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 source 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 was assessed using 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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:* One-way analysis of variance (ANOVA), simple linear regression, and multiple linear regression were used to compare mean intelligence grades by water fluoride levels and to assess the association between grades and urinary fluoride. Consideration of heterogeneity of variance (for ANOVA) was not reported. Regression diagnostics were not

used to test model assumptions for linearity, normality, and homogeneity. Given the ordinal nature of the intelligence grade variable (score from 1 to 5), ordinal logistic regression would have been a more appropriate method. There was no adjustment for area-level clustering in multivariate analyses (although subjects were selected via stratified cluster sampling from two areas). Although the analysis used individual-level exposures rather than area-level exposures, if the exposure levels within a certain area were highly correlated (which might be expected), then the results might still be biased. However, the overall impact on the effect estimates is expected to be minimal given the use of individual-level data and adjustment for important covariates.

        - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements and the consideration of key covariates, but it was limited by the cross-sectional study design and lack of addressing dilution in the urine samples.

E.1.9. Seraj et al. (2012)

E.1.9.1. 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 association between water fluoride and IQ score (adjusted β per 1-mg/L increase in water fluoride = −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.

E.1.9.2. 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, sex, 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, sex, dental fluorosis intensity, and educational levels (child's and parents') were evaluated as important covariates. Other covariates such as smoking were not discussed. Information was obtained from a detailed questionnaire. Lead was measured but found only 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. Based on water quality maps, co-exposure to arsenic is likely not a major concern in this area.
  - *Potentially important study-specific covariates:* Arsenic.
    - *Direction/magnitude of effect size:* Conceptually, if there were differential amounts of arsenic in the different villages, co-exposure to arsenic could bias the association, with the direction of the bias dependent on where the arsenic was present; however, arsenic was not expected to be a major concern in this study area based on water quality maps.
  - *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 covariates, 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 whether 21 children were excluded based on exclusion criteria or whether 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 minimal, 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 the 4000 UV-Vis spectrophotometer in the environmental health engineering laboratory of the Public Health School of the Tehran University of Medical Sciences. Specific details were not provided on methods of collection or sample locations or whether 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, and they did not 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 size:* The presence of dental fluorosis in all groups indicates that there may have been different exposures 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 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 to subjects' exposure status (++ for blinding). Overall rating for methods and blinding = +.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that outcomes were assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* All outcomes outlined in the abstract, introduction, and methods were reported. However, because the study author did not report the method for

obtaining the betas in Table 4 of the study, it is not clear whether these were adjusted or unadjusted regression coefficients.

- *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 methods for comparisons of IQ level by exposure groups were reasonable (ANOVA, post hoc test, and Kruskal-Wallis test), but consideration of heterogeneity of variance was not reported. Clustering at the village levels was not accounted for in multivariate analyses, which used area-level water fluoride levels. Because the exposure levels within a certain area are highly correlated (which might be expected), the results are likely to be biased. There was adjustment for some individual-level important covariates, and the children were from five rural areas with similar general demographic and geographic characteristics and were comparable in terms of SES and parental occupations. These factors are expected to mitigate some of the impact of lack of accounting for clustering, and the overall impact on the effect estimates is expected to be minimal.
    - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and outcome. Study strengths include addressing potential key covariates, but it was limited by the cross-sectional study design and the group-level exposure data.

E.1.10. Soto-Barreras et al. (2019)

E.1.10.1. 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: Results were not presented to evaluate an association between fluoride exposure and IQ but to compare fluoride levels within IQ grades. For this reason, the results of this study are not comparable to

other studies that evaluated IQ scores by fluoride exposure levels. No significant differences in measured fluoride levels across IQ grades were observed.

E.1.10.2. Risk of Bias

- Author contacts:
  - Authors were not contacted for additional information to inform the risk-of-bias evaluation 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 there were no specific details 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 have been similar to 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 covariates were considered when evaluating associations between fluoride exposure and intelligence; covariates were considered only when evaluating associations between fluoride levels and dental caries. According to Table 4 of the study, there was no significant association between IQ grade and age, sex, parental education, or SES status. No other information was reported or considered. There is no information on potential co-exposures. According to 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 covariates:* Arsenic.
    - *Direction/magnitude of effect size:* The impact on the direction and magnitude of effect size 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

Commented [A116]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

together in the water, it would likely bias the association away from the null; however, if they occur in different areas, there is potential to bias the association 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 size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.

- 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 Kolmogorov-Smirnov test was used to determine variable distribution. The Kruskal-Wallis test was used to compare exposure levels between IQ grades with Dunn's post hoc test. Multivariate logistic regression was used to estimate the association between presence of dental caries and various risk factors. Fluoride levels in drinking water and urine and fluoride exposure dose were compared across intellectual grades. Children were from 13 schools selected via stratified cluster sample design. There was no adjustment for clustering at the school level or for the sampling design. Although the analysis used individual-level exposures rather than area-level exposures, if the exposure levels within a certain school were highly correlated (which might be expected), then the results might still be biased. The large number of clusters (13 schools) makes clustering less of a concern, and the impact on the effect estimates is expected to be minimal.
    - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and blinding of outcome assessor to participants' fluoride exposure, but it is limited by the cross-sectional study design, lack of accounting for urine dilution, and lack of consideration for potential exposures to arsenic in the study area. Although the study is considered to have low potential for bias overall, the focus of the study was to evaluate the relationship between fluoride exposure and lower rates of dental caries. In terms of evaluating an association between fluoride exposure and IQ scores, the study is limited by the way the data were reported.

E.1.11. Sudhir et al. (2009)

E.1.11.1. 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: Significant increase in mean and distributions of IQ grades (i.e., increase in proportion of children with intellectual impairment) with increasing drinking water fluoride levels.

E.1.11.2. 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 lifelong 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 (rather, each had 243–267). Participation rates were 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 sex 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 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 non-fluoridated). SES was measured using the Kakkar socioeconomic 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 socioeconomic 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 covariates, 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 covariates:* Key covariates age, sex, and measures of SES were similar between exposure groups; however, arsenic was not considered. 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 size:* Conceptually, the presence of arsenic would potentially bias the association away from the null if present with fluoride. Deficiencies in iodine would likely bias the association away from the null if present in areas of high fluoride but toward the null if present in areas of non-high fluoride. Neither of these were considered issues in this study for reasons noted above.
- *Basis for rating:* Probably low risk of bias based on indirect evidence that the key covariates were considered, co-exposure to arsenic was 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 the 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

the 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 the criteria was given a bottle for water collection, and the next child was given a bottle for collection only if the water source was different from that of a previous child. Children were asked to collect a water sample from the source that was used in the initial 10 years of their life (and that sample was collected the next day). It was not reported whether 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 size:* 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:* Raven's standard progressive matrices (1992 edition) was used to assess IQ. Raven's test is a standard test; although there is no information provided to indicate that the methods were reliable and valid in this study population, the 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 email 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 assessors were not blind to participants' fluoride exposure 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:* Probably low risk of bias (+)
    - *Summary:*

- *Statistical analyses:* Chi-square test and Spearman rank correlation were used to assess the association between four different fluoride levels and IQ grades. Area-level exposures were used. Clustering of children within the four areas was not accounted for in the analysis; however, because multiple villages were included in each fluoride exposure level, clustering was less of a concern and the impact on the effect estimates was expected to be minimal.
- *Other potential concerns:* None identified.

o *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 of risk of bias identified.

- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include verification of exposure measurements and consideration of key covariates, but it was limited by the cross-sectional study design and lack of information on blinding during outcome assessment.

E.1.12. Till et al. (2020)

E.1.12.1. 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 and water fluoride concentration (with or without adjusting for maternal urine) in formula-fed or breastfed infants or fluoride intake from formula.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower performance IQ with water fluoride per 0.5-mg/L increase by breastfeeding status (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 per 0.5-mg/L increase in formula-fed children (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.

E.1.12.2. Risk of Bias

- Author contacts:
  - ~~Authors were not contacted for additional information because it was not necessary.~~Authors were not contacted for the 2020 publication. Authors were contacted in June 2019 for additional information on the Green et al. (2019) publication. Information obtained from that correspondence may have been used for additional information in the 2020 publication.

Commented [A117]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

- 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 = 1,945) or the subset without complete water fluoride and covariate data (n = 203). A table of characteristics of the study population was provided. Approximately half of the children lived in non-fluoridated 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 sex and age at testing, maternal education, maternal race, secondhand smoke in the home, and HOME score. City was considered but excluded from the models. Covariates 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 because it receives water mainly from municipal water supplies that monitor for lead and arsenic, and the lack of information is not considered to appreciably bias the results. In addition, a previous study on this population (Green et al. 2019) conducted sensitivity analyses on co-exposures to lead and arsenic. Results from these sensitivity analyses support the conclusion that co-exposures to lead and arsenic are not likely a major concern in this study population.
  - *Potentially important study-specific covariates:* All key covariates 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 covariates were considered 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 size:* There is not any specific direction or magnitude of bias expected. Urinary fluoride levels are reflective of recent exposure. The possibility of 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, which is considered a gold standard test. It is appropriate for both the study population and age group. 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 whether 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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:* 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 warrants exploration of city as a random effect rather than a fixed effect (as would be the case by having it included as a covariate). Even including individual-level maternal urinary fluoride might not fully account for lack of a city effect, 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 significantly impact the effect estimates.
    - *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 of risk of bias identified.
- ***Basis for classification as low 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 consideration of key covariates.

E.1.13. Trivedi et al. (2012)

E.1.13.1. 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 mean IQ score in the high fluoride villages (92.53 ± 3.13) compared with the low-fluoride villages (97.17 ± 2.54) in boys and girls combined (and by sex).

E.1.13.2. Risk of Bias

- Author contacts:
  - Authors were contacted in September of 2017 to obtain additional information for risk-of-bias evaluation. Additional information provided by the authors informed the rating decision for the following risk-of-bias domains: Selection, Attrition, Detection (exposure assessment), Detection (outcome assessment).
- Population selection:
  - *Rating*: Probably low risk of bias (+)
  - *Summary:* There is insufficient information provided on the sampling methods to determine whether 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 was 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 whether the differences in numbers were based on different participation rates or whether 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 whether 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 of 80% attendance was a limiting factor that resulted in 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.

Commented [A118]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

- 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 ages 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 arsenic and lead were not considered, the study authors provided physicochemical analyses for the water samples from the six different villages. While the authors did not specifically analyze lead or arsenic in the water samples, these physicochemical analyses suggest that differential lead or arsenic exposure was unlikely. Moreover, based on water quality maps, arsenic was not expected to be a major concern in this study area. According to the information from the water quality maps and the physiochemical analysis of the water provided, there is indirect evidence that neither arsenic nor lead were a concern in this study population.
  - *Potentially important study-specific covariates:* Key covariates age, sex, and measures of SES were similar between exposure groups; however, arsenic was not considered. 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 size:* Conceptually, the presence of arsenic would potentially bias the association away from the null if present with fluoride, or toward the null if present in the reference group; however, for reasons noted above, arsenic is not considered a concern in this study population.
  - *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 covariates 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 whether analyses were restricted to these children or whether 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 fluoride villages versus high fluoride 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 the second water sample was collected.
  - *Direction/magnitude of effect size:* 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 the impact on the direction and magnitude of the effect size 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.

- 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 a 97% reliability rate in relation to the Stanford-Binet Intelligence Scale (+ for methods). Blinding or other methods to reduce bias were 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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 high risk of bias (−)
  - *Summary:*
    - *Statistical analyses:* Mean IQ scores in low and high fluoride villages were compared using a t-test. Consideration of heterogeneity of variances was not reported. Results are reported as means and standard errors of the means, with p-values for significant differences. Area-level exposures were used. There was no accounting for clustering of children within the villages, and comparative analyses did not account for covariates. Urinary fluoride was not considered in the comparative analyses. The lack of individual exposure levels and the lack of accounting for clustering are likely to bias the standard error of the difference in mean IQ levels between the high- and low-fluoride villages and make the differences appear stronger than they actually are.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the statistical analyses did not account for clustering, and this lack of accounting could bias the association. There were no other potential threats of risk of bias identified.
- ***Basis for classification as low 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 covariates, but the study was limited by the cross-sectional study design. Another limitation was the lack of accounting for clustering, which may bias the standard error of the differences, making the effect appear stronger than it actually is; however, this does not change the nearly 5-point difference in IQ scores between the two villages.

E.1.14. Wang et al. (2012)

E.1.14.1. Study Details

- ***Study design:*** Cross-sectional
- ***Population:*** Children aged 8–13 years [possibly the same study population as Xiang et al. (2003a)]
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Province, 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 endemic versus non-endemic regions, as reported in Xiang et al. (2003a); when the high-exposure group was broken into four exposure groups based on fluoride intake, a dose-dependent decrease in IQ and increase in % with low IQ was observed;

significant correlation between total fluoride intake and IQ (r = −0.332); for IQ <80, adjusted OR of total fluoride intake per 1 mg/(person/day) was 1.106 (95% CI: 1.052, 1.163).

E.1.14.2. 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 have the same study population as Xiang et al. (2003a) and Xiang et al. (2011); however, it does not cite these studies as providing additional information, and the 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 whether study subjects were 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, 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 whether arsenic was also present in the drinking water. However, a publication by Xiang et al. (2013) in the same study areas indicates that Xinhuai (the low-fluoride area) had significantly higher arsenic levels compared with Wamiao (the endemic fluorosis area), which would bias the association toward the null. Areas were stated to be similar in annual per capita income, transportation, education, medical conditions, 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. 2003a; Xiang et al. 2011) indicate that iodine and lead are not concerns.
  - *Potentially important study-specific covariates:* 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 with the high fluoride area. Because there were significant effects on IQ observed in the high fluoride areas, the impact of co-exposure to arsenic is less of a concern. The presence of arsenic in the control village may cause an underestimation of the effect of fluoride, but despite this potential impact, a significant association between fluoride exposure and IQ was reported.

  - *Direction/magnitude of effect size:* Presence of arsenic in this study population would potentially bias the association toward the null.

- *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key covariates were considered, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency were not attributing to the association observed in this study. The potential for bias toward the null combined with the reported significant association increases confidence in the observed effect.

- 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 size:* There is potential for exposure misclassification based on calculating fluoride intake based on measurements from a few select subjects rather than all subjects. The potential impact on the direction and magnitude of effect size 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 total fluoride 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:* 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 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-blind throughout the testing process (++ for blinding). Overall rating = ++.
    - *Basis for rating:* Definitely low risk of bias based on direct evidence that the outcome was assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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:* Logistic regression analysis was used to determine the odds of having low IQ with increasing fluoride intake. Analyses and methods

are not well described. There is no mention of what tests were used for the mean IQ comparison by village; however, statistical software (SPSS) was used, suggesting appropriate tests were applied. Simple linear regression analyses were conducted to evaluate associations between total fluoride intake and children's IQ or low IQ rate. There is no evidence that regression diagnostics were used to test model assumptions for linearity, normality, and homogeneity. Clustering at the village level was not accounted for in the analyses. The overall impact of these factors on effect estimates is expected to be minimal given the use of individual-level data and adjustment for important covariates.

    - *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 of risk of bias identified.
- ***Basis for classification as low 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 individual measurements to calculate fluoride intake. All key covariates were accounted for in the study design or analysis, but there is potential for the presence of arsenic to bias the association toward the null.

E.1.15. Wang et al. (2020b)

E.1.15.1. 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 associations between IQ score and water fluoride (adjusted $\beta = -1.587$ per 1-mg/L increase) and urinary fluoride (adjusted $\beta = -1.214$ per 1-mg/L increase) in boys and girls combined based on both quartiles and continuous measures. No significant modification effect of sex.

E.1.15.2. 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 non-endemic 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 whether the participation differed by area. However, there was a 93% recruitment rate, and the 13 excluded due to missing data were 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 were not exposed to other neurotoxins such as lead, arsenic, or mercury nor were they iodine-deficient. Final models included age, sex, child's BMI, maternal and paternal education, household income, and low birth weight. The other covariates that were considered are unclear as the authors only noted that the covariates 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. Covariates such as parental BMI, behavioral and mental health disorders, IQ, and quantity and quality of the caregiving environment were not considered.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is direct evidence that the key covariates were considered and indirect evidence that the methods for collecting the information were valid and reliable and that co-exposure to arsenic was 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 of whether the urine samples accounted for dilution.
    - *Direction/magnitude of effect size:* Not accounting for dilution could cause some exposure misclassification. The impact on the direction and magnitude of effect size 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, 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 noted 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes 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 (+)

- <u>*Summary*</u>*:*
  - *Statistical analyses:* Logistic and multivariate regression models accounting for covariates were used. Results are presented as betas or ORs and 95% CIs. Regression diagnostics were conducted for all models, including examination of multicollinearity, heteroscedasticity, and influential observations. Mediation and interaction analyses were appropriate. There is no evidence that the stratified and multistage random sampling approach for subject selection was accounted for in the analyses by using sampling weights or accounting for clustering using random effect models; however, selected villages were similar in population and general demographic characteristics. Given the use of individual-level data and adjustment for important covariates, the impact on the regression coefficients is likely to be minimal.
  - *Other potential concerns:* None identified.
- <u>*Basis for rating*</u>*:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and no other potential threats of risk of bias were identified.

- ***Basis for classification as low 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 covariates were considered in the study design or analysis.

E.1.16. Xiang et al. (2003a)

E.1.16.1. Study Details

- ***Study design:*** Cross-sectional
- ***Population:*** Children aged 8–13 years
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Province, 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 (a non-endemic fluorosis area); additional breakdown of the Wamiao area into five water fluoride exposure groups.
- ***Reported association with fluoride exposure:*** Yes: Significantly lower IQ scores observed with water fluoride levels of 1.53 mg/L or higher. Percentage of subjects with IQ scores below 80 was significantly increased at water fluoride levels of 2.46 mg/L or higher. Significant inverse correlation between IQ and urinary fluoride ($r = -0.164$). Mean IQ scores for children in the non-endemic region (100.41 ± 13.21) were significantly higher than the endemic region (92.02 ± 13.00).

E.1.16.2. 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 in the study; 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/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 age, sex, 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 sex, age, family income, and parental education were considered. 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 with Wamiao (the endemic fluorosis area). This is likely to bias the association 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 covariates:* 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 with the high fluoride area. Because there were significant effects on IQ observed in the high fluoride areas, the impact

of co-exposure to arsenic is less of a concern. The presence of arsenic in the control village may cause an underestimation of the effect of fluoride, but despite this potential impact, there was still a significant association between fluoride exposure and IQ.

- *Direction/magnitude of effect size:* Presence of arsenic in this study population would potentially bias the association toward the null.

- *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key covariates were taken into account, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency were not attributing to the effect observed in this area. The potential for bias toward the null, combined with the reported significant association increases confidence in the observed effect.

- 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 non-endemic 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 size:* 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 likely bias the association 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.

- 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:* There is no mention of the tests conducted, but data were stated to be analyzed using SAS, suggesting appropriate tests were applied. Results provided in the tables indicate that t-tests comparing IQ values between the villages (overall and by sex) were conducted, but it was not reported that heterogeneity of variance was assessed. In addition, correlations between IQ and age, family income, and parents' education level were tested with Pearson's correlation. There is no evidence that a test for trend was conducted to evaluate the stated "significant inverse concentration-response relationship between the fluoride level in drinking water and the IQ of children."
    - A potential concern raised by the NASEM (2020) committee's review was the lack of accounting for relationships in exposure between persons from the same village. Given only two villages were included and the analyses consisted of village-level comparisons (no use of individual-level covariate data), it is likely that the standard error of the difference in mean IQ between fluoride in water exposure groups will be biased, making differences appear stronger than they actually are. Without controlling for village effects and given the large differences in fluoride concentrations and IQ levels between villages, the apparent dose-response relationship could be due to a village effect in addition to a fluoride effect. However, a dose-response relationship is apparent within the "exposed" village, diminishing the concern for a village-only effect and likely minimizing the impact on the effect estimates.
    - *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 of risk of bias.

- ***Basis for classification as low 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 blinding of outcome assessor to exposure but is limited by the cross-sectional study design and lack of accounting for urine dilution. All key covariates were considered in the study design or analysis, but there is potential for the presence of arsenic to bias the association toward the null.

E.1.17. Xiang et al. (2011)

E.1.17.1. Study Details

- ***Study design:*** Cross-sectional
- ***Population:*** Children aged 8–13 years [same study population as Xiang et al. (2003a)]
- ***Study area:*** Wamiao and Xinhuai villages located in Sihong County, Jiangsu Province, 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 linear trend across 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 observed at ≥0.05 mg/L serum fluoride.

E.1.17.2. 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 was the same as that used in the Xiang et al. (2003a) study, but a few more measurements were available and different analyses were conducted. The comparison population was considered the same 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 (Xiang et al. 2003a), 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 sex were adjusted for in the analysis. In the previous report, no significant associations were observed between groups for family income and parents' education (Xiang et al. 2003a). 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 and neither 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 with Wamiao (the endemic fluorosis area). This is likely to bias the association 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 covariates:* 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. Because there were significant effects on IQ observed in the high fluoride areas, the impact of co-exposure to arsenic is less of a concern. The presence of arsenic in the control village may cause an underestimation of the effect of fluoride, but despite this potential impact, there was still a significant association between fluoride exposure and IQ.
  - *Direction/magnitude of effect size:* Presence of arsenic in this study population would potentially bias the association toward the null.
- *Basis for rating:* Probably low of risk bias because there is indirect evidence that the key covariates were considered, methods used for collecting the information were valid and reliable, and co-exposures to arsenic and lead and iodine deficiency were not attributing to the effects observed in this area. The potential bias toward the null, combined with the reported significant association increases confidence in the observed effect.

- 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 the authors combined the lower two quartiles. After combining the two lower quartiles, the exposure levels ranged from <0.05 mg/L (Q1 + Q2) to >0.08 mg/L (Q4).

- *Direction/magnitude of effect size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- 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 conducted were appropriate for the study. Chi-square tests were used to compare categorical variables, and multiple 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. Although only two villages were included, in the analyses that consisted of village-level comparisons, it is likely that the standard error of the difference in mean IQ between villages is biased. This is less of a concern for the mean IQ comparisons across quartiles of serum fluoride levels and for the logistic regression analyses of risk of low IQ and individual-level serum fluoride levels. 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 and likely minimizing the impact on the effect estimates.

        - *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 of risk of bias identified.
- ***Basis for classification as low 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 covariates were considered in the study design or analysis, but there is potential for the presence of arsenic to bias the association toward the null.

E.1.18. Yu et al. (2018)

E.1.18.1. 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 in mean IQ scores in high water fluoride areas (>1.0 mg/L; 106.4 ± 12.3 IQ) compared to the normal water fluoride areas (≤1.0 mg/L; 107.4 ± 13.0). 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 a 4.29 decrease in IQ score (95% CI: −8.09, −0.48).

E.1.18.2. Risk of Bias

- Author contacts:
    - Authors were contacted in September 2018 to obtain additional information for the risk-of-bias evaluation. Additional information provided by the authors

Commented [A119]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

informed the rating decision for the following risk-of-bias domains: Detection (outcome assessment).

- 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 multistage 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 were 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 covariates:* All key covariates were considered in this study.

    - *Direction/magnitude of effect size:* 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 covariates and co-exposures were considered.
- 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 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 size:* 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 the 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 whether the 40 children randomly assigned to the classroom were specific to the village or whether 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 assessed using instruments that were valid and reliable, and that the outcome assessors were blind to participants' fluoride exposure.

- 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 used were appropriate for the study. Univariate and multivariable piecewise linear regression models were used to estimate the associations between water fluoride or urinary fluoride levels and IQ scores. Multiple logistic regression analysis was used to evaluate the association between water or urinary fluoride levels and IQ degree using the normal intelligence group as the control. Sensitivity analyses were conducted. There is no evidence that residual diagnostics were used to examine model assumptions or that the complex sampling design (stratified multistage random sampling) was accounted for in the analysis using sampling weights and adjustment for clustering. The impact of these factors on the effect estimates is expected to be minimal given the use of individual-level data and adjustment for numerous important covariates.
    - *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 of risk of bias identified.

- ***Basis for classification as low 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 covariates, including potential co-exposures, were considered in the study design or analysis.

E.1.19. Zhang et al. (2015b)

E.1.19.1. 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 children's serum fluoride ($r = -0.47$) and urinary fluoride ($r = -0.45$); significant difference in mean IQ score for high-fluoride area (defined as >1 mg/L in drinking water; 102.33 ± 13.46) compared with control area (<1 mg/L; 109.42 ± 13.30).

E.1.19.2. 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, and 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 were considered 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 age, sex, educational levels of parents, drinking water fluoride (mg/L), and levels of thyroid hormones (T3, T4, and TSH). Authors report that the study areas were not affected by other contaminants such as arsenic or iodine, and residents were of similar physical and mental health status. Other important covariates (maternal demographics, smoking, reproductive health) were not considered. Covariate data were obtained from a study questionnaire.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* 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 covariates, including potential co-exposures, were considered.
- 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 content in drinking water (W-F), serum (S-F), and urine (U-F) was 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 size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- 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:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Associations between serum and urinary fluoride levels and IQ score were estimated using general linear models and multivariate linear regression by COMT polymorphism. Normality (Kolmogorov-Smirnov test) was evaluated for all continuous variables. There is no evidence that residual diagnostics were used to examine model assumptions or that the complex sampling design (stratified multistage random sampling) was accounted for in the analysis using sampling weights and adjustment for clustering. The impact of these factors on the regression effect estimates is expected to be minimal given the use of individual-level data and adjustment for numerous covariates.
    - *Other potential concerns:* None identified.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that the statistical analyses were appropriate and there were no other potential threats of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, blinding of outcome assessor to participants' fluoride exposure, and consideration of key covariates including potential co-exposures.

E.2. Other Neurodevelopmental Studies

E.2.1. Barberio et al. (2017b)

E.2.1.1. 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 per 1-µmol/L increase (1.02; 95% CI: 1.00, 1.03) when Cycles 2 and 3 were combined. No significant associations with creatinine-adjusted or specific gravity-adjusted urinary fluoride. No significant association between urinary fluoride and ADHD.

E.2.1.2. 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 represented only 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 their current home for 3 or more years. Covariates such as parental behavioral and mental health disorders, smoking, and nutrition

were not discussed. The study used data from the Canadian Health Measures Survey, which consists of a nationally representative sample of Canadians. Most Canadians (~89%) receive water from municipal water supplies, which monitor for levels of lead and arsenic. Therefore, 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 covariates:* All key covariates were considered in this study.
  - *Direction/magnitude of effect size:* Not applicable.
- *Basis for rating:* Probably low risk of bias based on direct evidence that key covariates 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 size:* There is not any specific impact on the direction or magnitude of effect size 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 likely 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 was 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:* Logistic regression analyses, adjusted and unadjusted for covariates, examined the associations between fluoride exposure and diagnosis of learning disability. Analyses were performed for Cycle 2 only (urinary fluoride and type of learning disability diagnosis), Cycle 3 only (urinary fluoride, water fluoride, and learning disability diagnosis), and Cycles 2 and 3 combined. Analyses used survey weights and bootstrapped weights to ensure proper computation of variance estimates. Results are reported as unadjusted and adjusted ORs with 95% CIs.
    - *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 of risk of bias identified.

- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in confounding and exposure. Study strengths include individual exposure measurements and the consideration of key covariates, but was limited by the cross-sectional study design and insensitive outcome measures.

E.2.2. Bashash et al. (2017)

E.2.2.1. 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 had data for the general cognitive index (GCI).
- ***Data relevant to the review:*** Adjusted and unadjusted associations between GCI and maternal or child's urinary fluoride concentrations.
- ***Reported association with fluoride exposure:*** Yes: Significant association between maternal urinary fluoride and GCI score (adjusted β per 0.5 mg/L increase = −3.15; 95% CI: −5.42, −0.87). No significant associations with children's urinary fluoride.

E.2.2.2. Risk of Bias

- Author contacts:
  - Authors were contacted for additional information on whether clustering was addressed. The authors provided results from additional models with cohort as a random effect, which informed the rating decision for the following risk-of-bias domains: Other. ~~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 neurodevelopmental 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 information on smoking habits was not included. Study populations appear to be similar, but there may be some differences because subjects were selected from two different cohorts that were recruited during slightly different time periods.

Commented [A120]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

  - *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 for whom 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, sex, birth weight, birth order, age at testing, maternal marital status, smoking history, maternal age at delivery, maternal IQ, education, and cohort, with additional testing for children's urinary fluoride, mercury, lead, and calcium. Sensitivity analyses were additionally adjusted for HOME score. Covariates 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 is included in the water quality control program in Mexico City and is not considered a concern in this population.
  - *Potentially important study-specific covariates:* All key covariates were addressed.
    - *Direction/magnitude of effect size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that key covariates, including other potential co-exposures, were considered, and indirect evidence that the methods used to collect the information were valid and reliable and that arsenic was 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 size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- 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. Statistical tests of bivariate associations (using Chi-square tests for categorical variables and analysis of variance [ANOVA]) were used to compare the means of the outcomes or exposures within groups based on the distribution of each covariate. Generalized additive models (GAMs) were used to estimate the adjusted association between fluoride exposure and measures of children's intelligence. Residual diagnostics were used to examine model assumptions and identify any potentially influential observations. Results are reported as adjusted effects and 95% CIs. In sensitivity analyses, regression models accounted for clustering at the cohort level by using cohort as a fixed effect in the models. Although using cohort as a random effect would be more appropriate, using individual-level exposure data and accounting for numerous important covariates in the models likely captured the cohort effect.

Commented [A121]: Change made in response to the [redacted] Reviewer comment below; see DocA1_Monograph for detailed response:

Reviewer comment (DocA1_Monograph, page 5): The NTP response to NASEM's comments indicate they contacted investigators on specific issues, and that some study authors responded while others did not. Recommendation: NTP should provide an assessment of how non-response would affect the grading of these studies. This may be tied to the NASEM concern regarding transparency in the risk-of-bias assessment. [redacted] requests an appendix and language in the Discussion detailing which studies received a change to their risk-of-bias determination.

Additional models with cohort as a random effect were also subsequently made available via personal communication with the study authors and showed similar results to the main model.

- *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 of risk of bias identified.

- ***Basis for classification as low risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, blinding of outcome assessor to participants' fluoride exposure, and the prospective cohort study design.

E.2.3. Bashash et al. (2018)

E.2.3.1. 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 (per 0.5-mg/L increase) and Conners' Rating Scales-Revised (CRS-R) scores, including Cognitive Problems and Inattention Index (adjusted $\beta = 2.54$; 95% CI: 0.44, 4.63), DSM-IV Inattention Index (adjusted $\beta = 2.84$; 95% CI: 0.84, 4.84), DSM-IV ADHD Total Index (adjusted $\beta = 2.38$; 95% CI: 0.42, 4.34), and ADHD Index (adjusted $\beta = 2.47$; 95% CI: 0.43, 4.50).

E.2.3.2. 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 considered 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 were adjusted for maternal age at delivery, years of education, marital status, smoking history, gestational age at birth, age at outcome assessment, sex, birth order, SES, cohort, and calcium intervention. Arsenic is included in the water quality control program in Mexico City and is not considered a concern in this population.
  - *Potentially important study-specific covariates:* 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 size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias based on direct evidence that key covariates 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 were 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 Conners' Rating Scales-Revised, which has been validated for the evaluation of ADHD. Mothers completed the CRS-R at the same follow-up visit in which the child completed the CPT-II tests. All tests were applied under the supervision of an experienced psychologist (++ for methods). Use of only parent reports and not teacher reports was noted by the authors as a study limitation because there is considerable variation between the two sources in terms of identifying ADHD-associated behaviors. 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- 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:* Bivariate analyses included Chi-square tests for categorical variables and ANOVA for continuous outcomes. Appropriate univariate statistics and transformations were performed before bivariate analyses. Residuals from fully adjusted linear regressions were checked and suggested skewness. Gamma regression with an identity link was used to examine the adjusted association between prenatal fluoride and each neurobehavioral outcome (instead of using log transformation). Generalized additive models were used to visually examine potential non-linearity. Sensitivity analyses examined impact of other covariates. Diagnostics tests were used to assess violations of the model assumptions and to identify remaining influential observations. The Benjamini-Hochberg false discovery rate (FDR) procedure was used to correct for multiple testing.
        - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Definitely or probably low risk-of-bias ratings in confounding, exposure, and outcome. Study strengths include individual exposure measurements, blinding of outcome assessor to participants' fluoride exposure, and the prospective cohort study design.

E.2.4. Choi et al. (2015)

E.2.4.1. 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, 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:*** No: None of the outcomes were significantly associated with fluoride exposure.

E.2.4.2. 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 whether the 51

children represented all first-grade children from this area or whether 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. Covariates were 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. Covariates 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 have been a concern.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is direct evidence that the key covariates were considered and indirect evidence that co-exposure to arsenic was 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 totals only 43, but the text indicates 8 children did not have a Dean Index because permanent 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 standard methods were likely used because analyses 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 was collected at home the following morning, 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 commonly used 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 size:* 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 the tests 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 using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.

- 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 appropriate. Multiple regression models evaluated the associations between exposure indicators and test scores after adjusting for covariates. Specific regression models are not described or referenced, just stated to be "standard regression analysis with confounder adjustment." The distributions of fluoride concentrations in urine and water were skewed and were log10-transformed to approximate a Gaussian distribution (test not specified). Results were reported as adjusted effects and 95% Cis. There was no evidence that residual diagnostics were used to examine model assumptions; however, the impact on the effect estimates is expected to be minimal.
    - *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:* Probably low risk of bias based on indirect evidence that the statistical analyses were appropriate and there were no other potential threats of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in the confounding, exposure, and outcome domains. Study strengths include individual fluoride measurements with blinding at outcome assessment likely. All key covariates and many other covariates were considered in the study design or analysis.

E.2.5. Li et al. (2004) [translated in Li et al. 2008a]

E.2.5.1. 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 (38.28 ± 1.10) groups; 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).

E.2.5.2. 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. Participants 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 were 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 covariates were specifically considered 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 which specific characteristics were considered. There were no significant differences in infant sex, 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 covariates:* Key covariates, including age, sex, and measures of socioeconomic status (SES), were similar between exposure groups; however, arsenic was not considered. 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 size:* Conceptually, the presence of arsenic would potentially bias the association away from the null if it were present with fluoride. Deficiencies in iodine would potentially bias the association 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. Neither of these are considered a concern in this study for reasons detailed above.
  - *Basis for rating:* Probably low risk of bias based on indirect evidence that the key covariates were considered, co-exposure to arsenic was 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:* Although authors did not discuss why only 91 of the 273 neonates available were randomly selected, 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 size:* 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 association 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 a 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 using instruments that were valid and reliable in the study population, and that the outcome assessors were blind to participants' fluoride exposure.
- Selective Reporting:
  - *Rating:* Probably low risk of bias (+)
  - *Summary:* The study authors reported numerous outcomes in sufficient detail; however, because a list of outcomes tested was not provided, 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 are described only as a t-test. Consideration of heterogeneity of variance was not reported. Results are reported as mean and standard deviations of neurological scores. Maternal urinary fluoride levels were used only to compare exposures between exposed and control groups. Infants in the control group were from four villages, and those in the exposed group were from five villages within the same district. Infants were randomly selected before they were assigned to exposed or control groups. In the comparisons, there was no accounting for clustering at the village level. It is likely that the standard error of the difference in mean neurobehavioral assessment scores between the high fluoride group and control group will be biased, making differences appear stronger than they actually are. However, the use of multiple villages per exposure group is likely to mitigate some of the impact of this lack of accounting for clustering, and the overall impact on effect estimates is expected to be minimal.
- *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 whether they were randomly selected to obtain equal numbers in the high-fluoride and control groups.

o *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 of risk of bias.

- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in the confounding, exposure, and outcome 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 was likely not a major concern according to groundwater quality maps.

E.2.6. Riddell et al. (2019)

E.2.6.1. 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. In addition, a significant association between

ADHD diagnosis (adjusted OR = 1.21; 95% CI: 1.03, 1.42) or hyperactivity/inattentive symptoms (adjusted β = 0.11; 95% CI: 0.02, 0.58) and community water fluoridation status. No significant associations with urinary fluoride levels.

E.2.6.2. 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 available only 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 age at testing, sex, ethnicity, BMI, parents' education, total household income, exposure to cigarette smoke inside the home, and log-transformed concurrent blood lead levels. Covariates such as parental behavioral and mental health disorders, quantity and quality of caregiving environment, and co-exposure to arsenic were not discussed. The study used data from the Canadian Health Measures Survey, which consists of a nationally representative sample of Canadians. Most Canadians (~89%) receive water from municipal water supplies, which monitor for levels of arsenic. Therefore, co-exposure to arsenic is not likely an issue in this population. 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 of data for covariates, but it is likely the questionnaires from the Canadian Health Measures Survey, which are considered standardized and validated.
  - *Potentially important study-specific covariates:* All key covariates were considered in this study.
    - *Direction/magnitude of effect size:* Not applicable.
  - *Basis for rating:* Probably low risk of bias because there is indirect evidence that the key covariates were considered, co-exposure to arsenic was likely not an issue, and methods used for collecting the information were valid and reliable.
- Attrition:
  - *Rating:* Probably low risk of bias (+)

- *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 considering 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.
- *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:* **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 Santé Publique du Québec. 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 an LOD of 0.006 mg/L. Values below the LOD were imputed with LOD/square root(2). Of the 980 samples, 150 (15%) were below detection.

    Chlorinated 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 size:* There is not any specific impact on the direction or magnitude of effect size 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 due to tap water or chlorinated water fluoride status, since 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 used only 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. For children aged 6–11 years, ADHD diagnosis was provided by parents, but youths aged 12–17 years completed the questionnaire themselves. Cycle 2 asked "Do you have a learning disability?"; if the subject answered "yes," he/she was asked to specify the type (four options were available and described). In Cycle 3, parents were asked directly whether they had ADHD, and children 12 years and older were asked whether 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 a doctor's diagnosis). Both were measured in both cycles. Blinding is likely not 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 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:* Robust logistic regression was used to examine the association between fluoride exposure and ADHD diagnosis, adjusting for covariates. 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. Multicollinearity was evaluated using variance inflation factor (VIF) statistics. Outliers with high studentized residuals, high leverage, or large Cook's distance values were removed from all analyses with urinary fluoride. All regressions were tested for interactions between fluoride exposure and age and between fluoride exposure and sex. Sensitivity analyses were conducted to test the different

survey cycles. There is no mention of adjustment for the complex survey design using survey weights or bootstrapped weights to ensure appropriate calculation of the estimated variances; however, the overall impact on effect estimates is expected to be minimal.

- *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 of risk of bias identified.

- ***Basis for classification as low 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 key covariates, but was limited by the cross-sectional study design and insensitive outcome measures.

E.2.7. Rocha-Amador et al. (2009)

E.2.7.1. 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.

E.2.7.2. 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 with F and As contamination is 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, 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 group 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:* Covariates included blood lead (PbB), age, sex, 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 outcome; therefore, arsenic co-exposure is not considered a major concern in this study. 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 covariate 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 whether this would vary by fluoride or arsenic levels.
  - *Potentially important study-specific covariates:* SES.
    - *Direction/magnitude of effect size:* There are insufficient data to determine the impact on the magnitude or direction of effect size. The impact on the direction of the association would likely depend on the association between fluoride exposure and SES.
  - *Basis for rating:* Probably high risk of bias based on indirect evidence that the SES was not considered 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 for 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 size:* 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 likely 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 was 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants’ fluoride exposure.
- 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:* Statistical analyses used log-transformed exposure variables (although rationale was not provided). Crude and partial correlations were calculated to evaluate associations between serum fluoride levels and TOCF scores. There is no other description of the regression model, and regression diagnostics to evaluate model assumptions are not presented; however, the overall impact on effect estimates is expected to be minimal.

    - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Probably low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and blinding of outcome assessor to participants' fluoride exposure, but it is limited by the cross-sectional study design, lack of consideration of SES in the study population, co-exposure with arsenic, and use of spot samples in a small population.

E.2.8. Valdez Jimenez et al. (2017)

E.2.8.1. 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 aged 3 to 15 months and evaluated for associations with first and second trimester maternal urine fluoride.
- ***Reported association with fluoride exposure:*** Yes: Significant association between log10-mg/L maternal urinary fluoride and MDI score during first trimester (adjusted $\beta = -19.05$; SE = 8.9) and second trimester (adjusted $\beta = -19.34$; SE = 7.46). No association between maternal fluoride during any trimester and Psychomotor Developmental Index (PDI).

E.2.8.2. 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 whether 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 the authors 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 second and third 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 score 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 analyses were conducted on the other factors, including sex, prior to conducting multivariable regression models. Some important covariates were not considered, including parental mental health, IQ, smoking, and potential co-exposures. Water quality maps indicate a potential for arsenic to be present in the study area.
  - *Potentially important study-specific covariates:* Arsenic is a potential co-exposure in this area of Mexico.
    - *Direction/magnitude of effect size:* If arsenic were present as a co-exposure, it would likely bias the association 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 who participated only 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 was assessed through morning urine samples and water fluoride levels collected from the children's homes. Sampling methodology was 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 size:* 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 assessed using instruments that were valid and reliable in the study population, and that the outcome assessor was blind to participants' fluoride exposure.
- 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 displays only data for trimesters 1 and 2. Although third trimester data were collected, they were not reported, likely because they were available for only 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:* Probably low risk of bias (+)
  - *Summary:*
    - *Statistical analyses:* Statistical analyses used log10-transformed exposure variables. Normality, homoscedasticity, and linearity assumptions were tested and satisfied for MDI and PDI scores. Bivariate analyses included

correlations, t-tests, and ANOVA. Multiple linear regression models by the first and second trimester of pregnancy were used to evaluate the association between maternal fluoride exposure and MDI and PDI scores. The best-fit model was selected using a "stepwise method," and the best-fit line was evaluated using "the curve fitting method." It is not further specified or cited what these methods entailed. Best-fit or goodness-of-fit statistics are not reported. It is unclear how a best-fit model could be selected when the authors state that all models adjusted for the same set of covariates regardless of significance, and these covariates also appear in the final model—presumably the best-fit model. It is unlikely that a stepwise method would retain all those covariates unless they were forced in the model. Residual analysis was conducted to assess model validity; however, there is no description of the results of the residual analysis. Nonetheless, the impact on effect estimates is expected to be minimal.

        - *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 into consideration in other domains (***Selection, Confounding***).
    - *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 of risk of bias identified.
- ***Basis for classification as low risk-of-bias study overall:*** Definitely low risk-of-bias ratings in exposure and outcome. Study strengths include individual exposure measurements and blinding of outcome assessor to participants' fluoride exposure, but it is limited by the cross-sectional study design and lack of accounting for potential co-exposures to arsenic.

E.2.9. Wang et al. (2020a)

E.2.9.1. 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 association between psychosomatic problems and urinary fluoride (per 1-mg/L increase; adjusted $\beta = 4.01$ [95% CI: 2.74, 5.28]) and increased risk of a T-score >70 with urinary fluoride (per 1-

mg/L increase; adjusted OR = 1.97 [95% CI: 1.19, 3.27]). No significant associations with ADHD or other measures of learning disability.

E.2.9.2. 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 was a pilot study, and it was not explicitly stated whether all eligible subjects participated in the study. There is no information on participation rates or whether 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, sex, 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, sex, 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 covariates:* SES.
  - *Direction/magnitude of effect size:* The impact on the direction and magnitude of effect size are 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 covariates were considered, that the methods for collecting the information were valid and reliable, and that co-exposure to arsenic was 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 size:* Spot urine samples account for only recent exposure. Although this could cause 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 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 were 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 whether subjects were selected based on areas with endemic fluoride or whether 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 assessors were blind to the participants' fluoride exposure.

- Selective Reporting:
  - *Rating:* Definitely low risk of bias (++)
  - *Summary:* All outcomes 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:* Multiple linear regression models were used to assess the association between urinary fluoride exposure and each behavioral outcome. Logistic regression was used to assess the risk of behavioral problems (T-scores >70) due to fluoride exposure. Sensitivity analyses were performed, with models adjusting for combinations of age, BMI, sex, mother migrated, father migrated, and urinary creatinine levels. Regression diagnostics to evaluate model assumptions are not described; however, the overall impact on effect estimates is expected to be minimal.
    - *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 of risk of bias were identified.
- ***Basis for classification as low 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 covariates were considered in the study design or analysis.

Appendix F. Mechanistic Data from Animal Studies

Table of Contents

Figures

A number of animal studies were available that presented mechanistic data in several effect categories (see Figure F-1). Limiting the data to studies with at least one exposure at or below 20 ppm fluoride drinking water equivalents (gavage and dietary exposures were backcalculated into equivalent drinking water concentrations for comparison) still provided a sufficient number of studies for evaluation of several mechanistic endpoints while allowing for a more focused look at exposure levels most relevant to human exposures. The following sections summarize the mechanistic data by effect category. Although there is some evidence of consistency in mechanistic effects, overall these data are insufficient to increase confidence in the assessment of findings from human epidemiological studies.

	Dose Level	
Mechanism	**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

Figure F-1. Number of Animal Mechanistic Studies for Fluoride by Mechanistic Category and Exposure Level

An interactive version of Figure F-1 and additional study details are available at https://public.tableau.com/app/profile/ntp.visuals/viz/FluorideTableauDashboards/ReadMe. 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 is summarized in the "All" column.

F.1. Neurotransmitters

Neurotransmitter and biochemical changes in the brain and neurons were considered 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 F-2). 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 low risk-of-bias studies that assessed changes in cholinesterase or acetylcholine (n = 11 drinking water studies) (Adedara et al. 2017a; Akinrinade et al. 2015a; Baba et al. 2014; Chouhan et al. 2010; Gao et al. 2008a; Gao et al. 2009; Khan et al. 2017; Liu et al. 2010; Mesram et al. 2016; Nkpaa and Onyeso 2018; Sun et al. 2000 [translated in Sun et al. 2008]), with the majority reporting evidence of an effect that is considered inconsistent with the phenotypic outcome (see Quality Assessment of Individual

Studies section for methods on determining which studies pose low risk of bias). 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 low risk of bias (Adedara et al. 2017a; Baba et al. 2014; Gao et al. 2009; 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 four of the five studies found statistically significant decreases in cholinesterase or AChE below 10 ppm. The five 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 (Akinrinade et al. 2015a; Gao et al. 2008a) 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 statistically significant only 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 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 low 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 low risk-of-bias studies. There were only five studies that evaluated dopamine and/or metabolites (Banala et al. 2018; Chouhan et al. 2010; Reddy et al. 2014; Sudhakar and Reddy 2018; Tsunoda et al. 2005). Four of the studies observed decreases in dopamine levels in the brain with exposures of less than 20 ppm fluoride (Banala et al. 2018;

Chouhan et al. 2010; Reddy et al. 2014; 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.

F.2. Biochemistry (Brain/Neurons)

Similar to the 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 low risk-of-bias studies (see Figure F-2). 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 low 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 in which the same protein was evaluated, results were not always consistent. These data are insufficient to increase confidence or support a change to hazard conclusions.

F.3. Histopathology

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. Histopathology of the brain was evaluated in 31 studies with concentrations at or below 20 ppm fluoride, of which 15 were considered low risk-of-bias studies (Adedara et al. 2017b; Akinrinade et al. 2015a; Bhatnagar et al. 2002; Bhatnagar et al. 2011; Chouhan et al. 2010; Guner et al. 2016; Jia et al. 2019; Jiang et al. 2014; Lou et al. 2013; McPherson et al. 2018; Mesram et al. 2016; Nageshwar et al. 2018; Niu et al. 2018; Pulungan et al. 2016; Zhao et al. 2019). In all but one low risk-of-bias study [Pulungan et al. (2016); gavage], animals were exposed to fluoride via drinking water. All low risk-of-bias studies were conducted in rodents, and all but three were conducted in rats (Wistar [seven studies], Sprague-Dawley [four studies], Long-Evans hooded [one study]). Overall, the low 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 low 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 low 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 low 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 low risk-of-bias studies handled the brains before fixation was complete, which can produce artifacts that can resemble dead neurons (Nageshwar et al. 2018; Zhao et al. 2019). Fixation and brain removal details were inadequately described in the remaining low 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 low 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 exposure at or below 20 ppm were reported in three of four low 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 low risk-of-bias studies with lower confidence in the outcome assessment (Jiang et al. 2014; Nageshwar et al. 2018; Niu 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 (>PND 80). 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 low risk-of-bias drinking water studies with higher confidence in the outcome assessment (Akinrinade et al. 2015a; Bhatnagar et al. 2011; Chouhan et al. 2010) and in three of four low 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 was evaluated in only one low risk-of-bias study (Bhatnagar et al. 2011; Guner et al. 2016). Pulungan et al. (2016), one of two low 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 (the study administered sodium fluoride via gavage; the 5-mg/kg/day dose was considered equivalent to 15.3 ppm fluoride in drinking water), nor were any of the results statistically significant.

F.4. Oxidative Stress

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. Oxidative stress

in the brain was evaluated in 25 studies that examined concentrations at or below 20 ppm fluoride, of which 15 studies had low potential for bias (Adedara et al. 2017a; Adedara et al. 2017b; Akinrinade et al. 2015b; Bartos et al. 2018; Chouhan and Flora 2008; Chouhan et al. 2010; Gao et al. 2008b; Gao et al. 2009; Guner et al. 2016; Khan et al. 2017; Mesram et al. 2016; Nageshwar et al. 2018; Nkpaa and Onyeso 2018; Shan et al. 2004; Zhang et al. 2015a). All of the low 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) (Adedara et al. 2017a; Adedara et al. 2017b; Akinrinade et al. 2015b; Bartos et al. 2018; Gao et al. 2008b; Gao et al. 2009; Guner et al. 2016; Khan et al. 2017; Mesram et al. 2016; Nageshwar et al. 2018; Nkpaa and Onyeso 2018; Shan et al. 2004; Zhang et al. 2015a) 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 low 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; Bartos et al. 2018; Gao et al. 2008b; Gao et al. 2009; Guner et al. 2016; Khan et al. 2017; Mesram et al. 2016; Nageshwar et al. 2018; Nkpaa and Onyeso 2018). Decreases in antioxidant activity using measures of superoxide dismutase (SOD) activity were reported in seven of eight low risk-of-bias studies (Adedara et al. 2017a; Adedara et al. 2017b; Akinrinade et al. 2015b; Khan et al. 2017; Mesram et al. 2016; Nageshwar et al. 2018; Nkpaa and Onyeso 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; Khan et al. 2017; Mesram et al. 2016; Nageshwar et al. 2018; Nkpaa and Onyeso 2018). A decrease in total antioxidant capacity (T-AOC) as a measure of antioxidant activity was also consistently reported in two low 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 low risk-of-bias studies (Adedara et al. 2017b; Nkpaa and Onyeso 2018).

Relative to the above-mentioned studies, 2 of the 15 low 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 low 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 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).

F.5. Apoptosis/Cell Death

Seven low 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.

F.6. Inflammation

Five low 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 low risk-of-bias studies. These data are insufficient to increase confidence or support a change to hazard conclusions.

F.7. Thyroid

Seventeen studies were identified that evaluated potential effects of fluoride on the thyroid with concentrations at or below 20 ppm (see Figure F-1). 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.

Mechanism	Mechanism Subcategory	Direction of Effect ↑	Direction of Effect ↓	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**

Figure F-2. Number of Low 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

An interactive version of Figure F-2 and additional study details are available at https://public.tableau.com/app/profile/ntp.visuals/viz/FluorideTableauDashboards/ReadMe. This figure displays study counts for low risk-of-bias studies, as these counts are most relevant to the text in this section. Counts for high risk-of bias studies or all studies combined can be accessed in the interactive figure. 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 be counted only once in the Grand Total column). Endpoints, species, strain, sex, and exposure duration are available for each study in the interactive figure.

Appendix G. 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 29, 2019	**Revised protocol:** Revised review protocol posted
September 16, 2020	**Revised protocol:** Revised review protocol posted

Appendix H. Supplemental Files

The following supplemental files are available at https://ntp.niehs.nih.gov/go/785076 or as *Sup01_Monograph_NASEM_Feb_2021.pdf.*

H.1. Protocol

NTP Protocol for Systematic Review of Human, Animal, and Mechanistic Evidence - Second Revision (September 16, 2020)
ntpprotocol_revised20200916_508.pdf

NTP Protocol for Systematic Review of Human, Animal, and Mechanistic Evidence - First Revision (May 29, 2019)
protocol_fluoridemay2019_508.pdf

NTP Protocol for Systematic Review of Human, Animal, and Mechanistic Evidence (June 2017)
protocol_fluoridejune2017_508.pdf

H.2. Response to NASEM Committee Letter Report

NASEM Committee Letter Report and Response for Monograph Only
Sup01_Monograph_NASEM_Feb_2021.pdf