August 2026
EPA-822-D-26-001

Protocol for Developing the Fluoride Human Health Toxicity Assessment

Protocol for Developing the Fluoride Human Health Toxicity Assessment

Prepared by:

U.S. Environmental Protection Agency
Office of Water (4304T)
Drinking Water Science and Engineering Division
Washington, DC 20460

EPA Document Number: EPA-822-D-26-001

August 2026

Disclaimer

This document is for informational purposes only. It does not represent and should not be construed to represent any agency determination or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Acknowledgments

This document was prepared by the Drinking Water Science and Engineering Division (DWSED), Office of Ground Water and Drinking Water (OGWDW), Office of Water (OW) of the U.S. Environmental Protection Agency (EPA). The DWSED/OGWDW/OW authors of this protocol document include Laura Dishaw (protocol lead) and Carlye Austin (protocol co-lead), Elizabeth Radke (lead of the Fluoride Human Health Toxicity Assessment), Thomas Bateson (co-lead of the Fluoride Human Health Toxicity Assessment), Todd Zurlinden (co-lead of the Fluoride Human Health Toxicity Assessment), Michael Dzierlenga (lead for the response to public comments on the assessment plan), Krista Christensen (no longer at the EPA), Paige Bommarito, Rachel Shaffer, Czarina Cooper, Lindsey Foote, Brittany Jacobs, Alexandra Lee, Casey Lindberg, Kelsey Miller, Hongyu Ru, Marie Russell, Paul Schlosser, Avanti Shirke, Anthony Su, J. Michael Wright, and Erin Yost. The agency appreciates the important contributions from EPA scientists in the Office of Applied Science and Environmental Solutions (OASES), including Chris Brinkerhoff, Sarah Gallagher, Rebecca Nachman, and Deborah Segal. The agency is also thankful for the valuable management oversight and review provided by Colleen Flaherty, Susan Euling, and Terra Haxton.

The systematic literature search and screening included in this document was prepared in collaboration with ICF under the EPA Contract 68HERC23D0003. ICF contributors included Samantha Snow, Meredith Clemons, Sorina Eftim, Alexander Lindahl, Wren Tracy, and Janielle Vidal.

Contents

Tables

Figures

Acronyms and Abbreviations

ADHD	attention deficit hyperactivity disorder
ADME	absorption, distribution, metabolism, and excretion
AI	artificial intelligence
AIC	Akaike information criterion
ATSDR	Agency for Toxic Substances and Disease Registry
AUC	area under the concentration curve
BIC	Bayesian Information Criteria
BMD	benchmark dose
BMDL	benchmark dose lower confidence limit
BMDS	Benchmark Dose Software
BMR	benchmark response
CASRN	Chemical Abstracts Service registry number
CDC	Centers for Disease Control and Prevention
CI	confidence interval
C_{max}	maximum blood concentration
C_{ss}	blood concentration at steady state
DoCTER	Document Classification and Topic Extraction Resource
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTXSID	Distributed Structure-Searchable Toxicity Substance Identifier
EFSA	European Food Safety Authority
EPA	U.S. Environmental Protection Agency
ER	extra risk
FDA	U.S. Food and Drug Administration
FRI	Fluorosis Risk Index
HAWC	Health Assessment Workspace Collaborative
HC	Health Canada
HED	human equivalent dose
HERO	Health and Environmental Research Online
HHS	U.S. Department of Health and Human Services
IARC	International Agency for Research on Cancer
IPCS	International Programme on Chemical Safety
IRIS	Integrated Risk Information System
IQ	intelligence quotient
LOAEL	lowest-observed-adverse-effect level
LOD	limit of detection
MCL	maximum contaminant level
MCLG	maximum contaminant level goal
MeSH	Medical Subject Headings
MRL	minimal risk level
NASEM	National Academies of Sciences, Engineering, and Medicine
NIEHS	National Institute of Environmental Health Sciences
NOAEL	no-observed-adverse-effect level
NPDWR	National Primary Drinking Water Regulation

NRC	National Research Council
NTP	National Toxicology Program
OASES	Office of Applied Science and Environmental Solutions
OECD	Organisation for Economic Co-operation and Development
OEHHA	Office of Environmental Health Hazard Assessment
OGWDW	Office of Ground Water and Drinking Water
OHAT	Office of Health Assessment and Translation
OMB	Office of Management and Budget
OPP	Office of Pesticide Programs
ORD	Office of Research and Development
OW	Office of Water
PBPK	physiologically based pharmacokinetic
PECO	populations, exposures, comparators, and outcomes
PET	positron emission tomography
PFAS	per- and polyfluoroalkyl substances
PK	pharmacokinetic
POD	point of departure
POD_{ADI}	average-daily-intake-based POD
QA/QC	quality assurance/quality control
REL	reference exposure level
RfD	reference dose
SEM	systematic evidence map
SDWA	Safe Drinking Water Act
SYR	Six-Year Review
TDI	tolerable daily intake
TFI	Thylstrup-Fejerskov Index
T_{max}	time to reach maximum blood concentration
TSIF	Tooth Surface Index of Fluorosis
UF	uncertainty factor
UF_A	animal-to-human uncertainty factor
UF_C	composite uncertainty factor
UF_D	database deficiencies uncertainty factor
UF_H	human variation uncertainty factor
UF_L	LOAEL-to-NOAEL uncertainty factor
UF_S	subchronic-to-chronic uncertainty factor
UL	tolerable upper intake level
UNII	Unique Ingredient Identifier
WHO	World Health Organization
WOS	Web of Science

Executive Summary

The U.S. Environmental Protection Agency (EPA) is committed to ensuring clean and safe drinking water for all Americans with the goal of protecting human health, including children's health. The EPA Office of Water (OW) is developing a new Fluoride Human Health Toxicity Assessment (*toxicity assessment* herein) that will review the available scientific information on the health risks of fluoride exposure and determine the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (i.e., toxicity value).

This Protocol for the Fluoride Human Health Toxicity Assessment (*protocol* herein) is the second step necessary for developing the toxicity assessment (Figure 1). The first planning step was the Preliminary Assessment Plan and Literature Survey for the Fluoride Human Health Toxicity Assessment (*assessment plan* herein) (U.S. EPA, 2026) that was released in January 2026. The protocol is an expanded version of the assessment plan. The protocol includes all of the content of the assessment plan (revised and updated to reflect feedback from the public) and additionally describes the methods that will be used to develop the toxicity assessment. The protocol is a methods document, not a toxicity assessment; it does not present the levels of fluoride exposure associated with harmful health effects. Determining these levels is part of the forthcoming draft toxicity assessment that will be developed according to the methods outlined in this protocol. The draft toxicity assessment will be released for public comment.

Toxicity assessments focus on determining the levels of exposure associated with potential harmful health effects of a contaminant to derive toxicity values that will protect against these harmful effects. Therefore, the Fluoride Human Health Toxicity Assessment will not consider beneficial effects of fluoride such as dental cavity prevention. Once final, the toxicity assessment can inform future steps under the Safe Drinking Water Act (SDWA). Although the EPA does not make recommendations about adding fluoride to drinking water, the final toxicity assessment could be used by the Centers for Disease Control and Prevention (CDC) to inform recommendations about the addition of fluoride to drinking water (United States Executive Office of the President Donald Trump, 2025).

Background. The EPA last evaluated the potential harmful health effects associated with fluoride exposure in 2010 (U.S. EPA, 2010a). In the EPA OW's 2010 assessment, the EPA determined that severe dental fluorosis, characterized by abnormal tooth enamel development in children, was the most sensitive harmful human health effect (i.e., occurring at the lowest fluoride exposure levels), and therefore, the toxicity value for fluoride derived in that assessment was based on this harmful health effect. In the fourth iteration of the EPA's process for reviewing existing drinking water regulations (Six-Year Review 4) (U.S. EPA, 2023a), the EPA concluded that the National Primary Drinking Water Regulation for fluoride was not a candidate for revision due to the need to evaluate emerging research on the association between fluoride exposure and neurodevelopment and cognition in children, which was under review by the National Toxicology Program (NTP) at that time. In its Six-Year Review 4 report, the EPA committed to considering the NTP (2024) report on fluoride exposure and neurodevelopment and cognition, when final, to inform development of an updated Fluoride Human Health Toxicity Assessment. The report by NTP was finalized in 2024 (NTP, 2024).

Following through on this commitment, the EPA is developing an updated Fluoride Human Health Toxicity Assessment, the objective of which is to determine the fluoride level that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (known as a toxicity value—specifically, noncancer reference dose or RfD).

Protocol Objective. The objective of this protocol is to describe the approaches and the peer-reviewed human health risk assessment methodology that will be used to develop the toxicity assessment. The protocol outlines the scope of the toxicity assessment, including how key science issues related to fluoride exposure and health effects evaluations will be addressed. It includes some elements (Sections 2 to 4) that were previously described in the assessment plan and have been revised in response to public comment. The protocol also describes the systematic review methods and fluoride-specific considerations that will be used to identify and evaluate evidence and derive a toxicity value for fluoride. The methods described in the protocol include how the EPA will: 1) prioritize the most informative studies, using a combination of automated prioritization processes and manual screening; 2) evaluate fluoride-specific studies to identify the best available science; and 3) identify studies that will be most useful for dose-response analysis and toxicity value derivation.

Consideration of Public Comments. The EPA publicly released the assessment plan (U.S. EPA, 2026), which describes the information that will be included in the Fluoride Human Health Toxicity Assessment, for a 30-day public comment period. The EPA reviewed all public input received (approximately 1,000 comments), considered all comments, and incorporated relevant feedback into protocol development. A summary of the public comments that the EPA received on the assessment plan and the agency's responses, including descriptions of how public comments informed protocol development, are provided in Appendix A.

Toxicity Assessment Scope. The protocol outlines the scoping and problem formulation for the Fluoride Human Health Toxicity Assessment that were described in the assessment plan, which have been revised for clarity in response to public comments. The scoping and problem formulation steps are intended to ensure the development of a focused, fit-for-purpose toxicity assessment. To help define scope, the EPA critically reviewed existing fluoride health effects assessments published by national and international agencies to evaluate their scientific conclusions about the types of harmful health effects associated with fluoride exposure. From this review, two well-established harmful health effects in children—dental fluorosis (i.e., abnormal tooth enamel development) and neurodevelopmental effects (i.e., abnormal brain development)—were identified as the most sensitive (i.e., occurring at the lowest exposure levels) harmful health effects associated with fluoride exposure (EFSA Scientific Committee, 2025; Health Canada, 2024; NTP, 2024; U.S. EPA, 2010b). The EPA has accepted these conclusions and will not repeat previous work to re-establish these hazards. However, this acceptance of hazard does not imply that the harmful effects occur at all doses or exposure levels. The EPA's future Fluoride Human Health Toxicity Assessment will independently evaluate the health effects data and develop a dose-response analysis (i.e., an analysis of how harmful health effects (response) change with increasing dose or exposure). The dose-response analysis will determine the doses or exposure levels at which harmful effects occur. Many

human studies on fluoride are available; therefore, the EPA's toxicity assessment will focus only on human data.

The toxicity assessment will focus on systematically identifying the best available human studies on dental fluorosis and neurodevelopmental effects of fluoride exposure to derive a toxicity value. For toxicity value derivation, estimates of fluoride exposure reported in human studies (e.g., in drinking water, urine) need to be converted into values that represent a person's daily fluoride intake to account for fluoride exposure from all potential sources, not just from drinking water. To determine daily fluoride intake, the EPA will ensure that fluoride exposure from sources other than drinking water is accounted for when examining toxicity values based on the concentration of fluoride in drinking water. Additionally, information on how fluoride moves through and is eliminated from the body will be incorporated into the toxicity value derivation process.

Literature Survey Results. As reported in the assessment plan, the EPA conducted a comprehensive, initial literature survey by conducting an initial search of the available published literature for studies on fluoride and reviewing the search results. Preliminary results of the survey were reported in the assessment plan, and full results are reported in this protocol. A total of 268,093 studies were identified. Machine learning-based artificial intelligence and other automated tools, in combination with manual review by scientific experts, were used to efficiently review the literature search results and identify studies most likely to be relevant to the toxicity assessment. Quality control checks were performed to ensure that the automated screening tools captured relevant studies. The studies prioritized through these methods then underwent further manual review by experts. The EPA identified 671 human studies examining fluoride exposure and dental fluorosis or adverse neurodevelopmental effects (e.g., cognition, behavior, and sensory/motor effects). The results of the preliminary literature survey are available for review via an interactive online application that allows users to search for, sort, and track the studies through each step of the survey. As outlined in this protocol, an updated literature search will be performed during toxicity assessment development to ensure that relevant studies published since the initial literature search can be incorporated into the toxicity assessment (Section 4.3).

Toxicity Assessment Methods. The protocol describes methods for conducting the steps of the future Fluoride Human Health Toxicity Assessment, including literature searching and screening; evaluation of studies and extraction of key information from them; dose-response analysis; and toxicity value derivation. The methods described in the protocol are consistent with agency guidance and peer-reviewed agency systematic review and human health risk assessment methodologies (U.S. EPA, 2022b, 2014a, b, 2012, 2011b, 2005, 2002b, 1998, 1996a, 1991). These methods identify the best available science in accordance with sound and objective scientific practices (SDWA 1412(b)(3)(A)(i) (U.S. EPA, 2019d)). The EPA's systematic review methods ensure that agency assessments are transparent, reproducible, and uphold scientific integrity. They promote the use of the best available, unbiased, peer-reviewed studies. The methods are consistent with Gold Standard Science as defined in (United States Executive Office of the President Donald Trump, 2025).

Approaches specific to the Fluoride Human Health Toxicity Assessment are described throughout the protocol, including planned methods to address the key science issues described in Section 2.4. The objective of the toxicity assessment and specific steps the EPA will follow to accomplish the objective are outlined in Section 3. The literature search, prioritization, and screening methods (Section 4) are customized to the challenges involved in managing the large fluoride health effects database and the toxicity assessment scope. The refined evaluation plan (Section 5) describes how reviewers will identify the studies most likely to be relevant for dose-response analysis based on considerations specific to the future Fluoride Human Health Toxicity Assessment. For study evaluation of epidemiology studies (Section 6.2), fluoride- and outcome-specific criteria are provided to supplement the generic study evaluation considerations that are relevant to all exposures and outcomes. These criteria include explicit descriptions of how different exposure metrics (e.g., drinking water, biomarkers) are interpreted during study evaluation (relevant to key science issues on validity of exposure measures and completeness of exposure characterization). Fluoride-specific pharmacokinetic models and approaches for evaluating them are provided in Section 6.3 (relevant to key science issue on understanding pharmacokinetics of fluoride in the body). The approach to data extraction is described in Section 7. Lastly, Section 8 describes methods for dose-response analysis, including selection of studies and analytical methods, with fluoride-specific considerations throughout (relevant to key science issue of dental fluorosis adversity).

Next Steps. The EPA will develop a draft toxicity assessment according to the methods described in this protocol. The toxicity assessment will summarize the harmful health effects associated with fluoride exposure during pregnancy and childhood and will determine the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (the RfD). The draft toxicity assessment will be released for external peer review and public comment (Figure 1). The EPA will consider the external peer review and public comments, revise the draft toxicity assessment as appropriate, and then publish a final toxicity assessment.

Figure 1. Steps in the Development of the Final Fluoride Human Health Toxicity Assessment

1. Introduction

Section Overview: The EPA is developing a fit-for-purpose toxicity assessment for fluoride. The toxicity assessment will determine the levels of fluoride exposure associated with harmful human health effects and will derive a toxicity value called an oral RfD. An RfD is an estimate of the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects. The oral RfD for fluoride may be used to inform future decisions about potential revisions to the existing fluoride drinking water standard under the Safe Drinking Water Act.

This document is the protocol, which is a roadmap for how the toxicity assessment will be conducted. The protocol includes the scope of the Fluoride Human Health Toxicity Assessment, which focuses on two harmful health effects in children related to fluoride exposure: dental fluorosis (which causes weakened tooth enamel and reduced tooth health) and effects on brain development (such as lower IQ). The agency will use a rigorous, transparent, systematic process to identify and evaluate human studies that assess the relationship between fluoride exposure and dental fluorosis or brain development. This process allows the EPA to select the best available data to establish an oral RfD for fluoride.

The U.S. Environmental Protection Agency (EPA) Office of Water (OW) is developing an updated fluoride human health toxicity assessment (*toxicity assessment* herein) based on a systematic review of studies on harmful health effects associated with fluoride exposure (U.S. EPA, 2025b). The purpose of the toxicity assessment is to derive a final toxicity value for oral fluoride exposure, called a reference dose (RfD). An RfD is an estimate of the amount of a chemical (in this case, fluoride) that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects.

The first planning step in developing the toxicity assessment was publishing the Preliminary Assessment Plan and Literature Survey for the Fluoride Human Health Toxicity Assessment (*assessment plan* herein) (U.S. EPA, 2026) in January 2026 for a 30-day public comment period. The assessment plan described the EPA's preliminary plan for developing the toxicity assessment, including the justification for conducting the toxicity assessment (background); intended scope or focus of the toxicity assessment (problem formulation); areas of scientific complexity (key science issues); toxicity assessment objectives (specific aims); the criteria that the EPA used to identify relevant data during the preliminary literature survey (literature search and screening); and results of the preliminary literature survey.

This Protocol for the Fluoride Human Health Toxicity Assessment (*protocol* herein) is an expanded version of the assessment plan and further describes how the toxicity assessment will be conducted. The protocol includes an overview of the assessment plan content (including revisions made in response to public comments) and additionally describes the methods for conducting the systematic review and dose-response analysis for the toxicity assessment. In developing the protocol, the EPA reviewed and considered all public comments received on the assessment plan during the 30-day public comment period (Appendix A).

To define the focus of the toxicity assessment, the EPA performed a critical review of peer-reviewed assessments of fluoride published by national and international agencies, including the National Institute of Environmental Health Sciences (NIEHS) National Toxicology Program (NTP) (NTP, 2024), the European Food Safety Authority (EFSA) (EFSA Scientific Committee, 2025), Health Canada (Health Canada, 2024), and the EPA (U.S. EPA, 2010a). The EPA's review of these assessments found that two adverse health outcomes in children—dental fluorosis (i.e., a condition of the tooth enamel that, when severe, can result in adverse effects on tooth health) and neurodevelopmental effects (e.g., decreased intelligence quotient [IQ])—were consistently identified as well-established and sensitive harmful health effects associated with fluoride exposure. These effects typically occur at lower exposure levels compared with other potential harmful health effects of fluoride, such as skeletal fluorosis. Therefore, a toxicity value based on dental fluorosis or neurodevelopmental effects is anticipated to protect against other adverse effects as well, which typically occur at relatively higher exposure levels. The toxicity assessment will focus on only human (epidemiological) studies because robust human studies are available for these outcomes (Section 2.3) and toxicity value derivation from human data eliminates uncertainties associated with extrapolation from animals to humans. The use of human data is preferred over animal data when robust human studies are available (U.S. EPA, 2022b). The proposed scope of the toxicity assessment underwent public comment as part of the assessment plan. This protocol includes the final scope of the toxicity assessment, which was revised for clarity in response to public comments.

The toxicity assessment will focus on dose-response analysis and derivation of a draft oral toxicity value for fluoride. It will not be a risk assessment and will not include an exposure assessment or a risk characterization (Figure 1-1). Additionally, it will not address the legal, policy, social, economic, or technical considerations involved in risk management. Once final, the toxicity assessment can be used by the EPA, states, Tribes, and local communities, along with specific exposure and other relevant information, to assess (under the appropriate regulations and statutes) potential risks associated with human exposure to fluoride. Specifically, the fit-for-purpose toxicity assessment will be used to inform future decisions about potential revisions to the existing fluoride drinking water regulation under the Safe Drinking Water Act (SDWA).

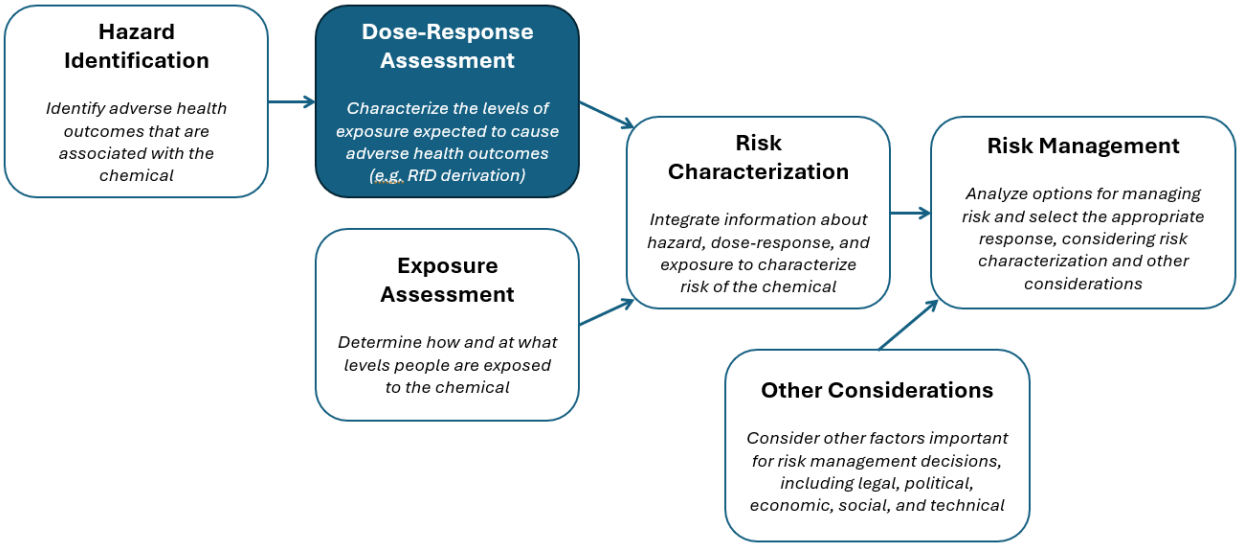


Figure 1-1. Overview of the Steps Involved in Risk Assessment and Risk Management

The focus of the Fluoride Human Health Toxicity Assessment will be dose-response analysis (shaded box).

As described in this protocol, the agency is implementing a robust systematic review process, based on the EPA's externally peer-reviewed and published human health risk assessment methodology (U.S. EPA, 2022b), to develop the toxicity assessment. Systematic review is a structured and documented process to identify, collect, and critically evaluate studies and data using explicit, prespecified methods. This approach is consistent with the SDWA requirements for the use of science in agency decision-making that direct the EPA "use (i) the best available, peer-reviewed science and supporting studies conducted in accordance with sound and objective scientific practices; and (ii) data collected by accepted methods or best available methods (if the reliability of the method and the nature of the decision justifies use of the data)" (SDWA 1412(b)(3)(A)(i)). These methods also ensure compliance with EPA quality assurance policies and procedures as described below and align with the Gold Standard Science objectives defined in (United States Executive Office of the President Donald Trump, 2025) by enhancing consistency, transparency, and reproducibility; reducing bias; and upholding scientific integrity. The systematic review process ensures critical review of uncertainties and findings of individual studies as well as the overall body of information. It also requires collaboration and input from many subject matter experts, which further reduces the potential for bias and error.

1.1. Quality Assurance

The environmental information operations described in this document were conducted in accordance with an approved EPA Quality Assurance Project Plan, applicable Agency Quality Policy Procedures, and CIO 2105 requirements. A summary of quality assurance (QA) processes is provided here; additional detail is available in Appendix B. As part of the assessment plan, the EPA conducted a comprehensive, preliminary literature survey to identify peer-reviewed scientific studies on associations between fluoride exposure and adverse dental fluorosis or neurodevelopmental effects. The methods used for the literature search and relevancy screening were initially described in the January 2026 Preliminary Assessment Plan and Literature Survey

(U.S. EPA, 2026). In this protocol document, the methods are described in Section 4 and Appendix C and were implemented in accordance with EPA QA policies and procedures [Quality Policy Procedures (U.S. EPA, 2008a) and CIO 2105.0 (formerly 5360.1 A2) (U.S. EPA, 2000)]. In the future during toxicity assessment development, relevant human studies identified in the preliminary literature survey will be critically evaluated for any concerns related to risk of bias or sensitivity (study evaluation). As part of future assessment development steps, the best available studies will then undergo data extraction, and those with datasets most amenable to dose-response analysis will be considered for RfD derivation. The EPA will also identify and evaluate studies on how fluoride moves through the body (pharmacokinetic information or models) and use that information to inform RfD derivation. The methods for these steps are described in this protocol. This protocol document is a planning document; it does not include any conclusions about the health risks from fluoride exposure or results from study evaluation or dose-response analysis. The protocol document describes the approaches and the peer-reviewed human health risk assessment methodology that will be used to develop the future human health toxicity assessment.

2. Scoping and Initial Problem Formulation

Section Overview: Scoping and problem formulation steps describe what the toxicity assessment will cover, the general approach that will be used to develop the toxicity assessment, and any important scientific issues that will need to be addressed by EPA scientists during the development of the toxicity assessment. This section begins with a high-level overview of how people are exposed to fluoride, potential human health effects of fluoride exposure, and how the EPA regulates fluoride in drinking water under the Safe Drinking Water Act (Section 2.1). Section 2.2 describes how the EPA's critical scientific review of published assessments on fluoride developed by national and international agencies informed the topics that the future Fluoride Human Health Toxicity Assessment will cover (scope, described in Section 2.3) and helped identify important scientific questions or areas of uncertainty that the toxicity assessment will address (key science issues, described in Section 2.4). Based on the best available science, the toxicity assessment will focus on two specific harmful effects in children—dental fluorosis and neurodevelopmental effects.

2.1. Background

To provide context for the development of the toxicity assessment, this section provides a brief overview of how humans can be exposed to fluoride as well as the history of fluoride regulation under SDWA. This focused overview is not intended to provide a comprehensive description of the available information on these topics.

2.1.1. Sources of Fluoride Exposure

Fluorides often refer to chemical compounds that contain the element fluorine (F). Fluorine is a highly reactive gas that does not occur in its pure elemental state in nature; instead, it primarily exists in compounds or as an ion. Examples of fluoride compounds include calcium fluoride, potassium fluoride, and sodium fluoride. The term "fluoride" can also be used to refer specifically to the fluoride ion (F^-). This protocol and the future toxicity assessment will be limited to the ionic form of fluoride (F^-) and related forms of fluoride that are expected to dissociate to the fluoride ion in water.

Human exposure to fluoride can occur from both natural and human-generated sources. Fluorine accounts for approximately 0.09 percent of the Earth's crust, making it the 13th most common element within the crust. Weathering of and leaching from fluorine-containing rock can lead to fluoride in surface water, groundwater, and soil. The composition and mineral content of underlying rock formations determine the subsequent levels of naturally occurring fluoride (predominantly F^-) in surface water and groundwater. Some fluoride-containing industrial chemicals (i.e., those that can dissociate to F^- in water) and the addition of fluoride to drinking water are two sources of fluoride in water that originate from human activities (WHO, 2004; ATSDR, 2003). Many organic fluoride-containing compounds, including many per- and polyfluoroalkyl substances (PFAS), do not readily break down to release ionic fluoride, due to the high stability of carbon-fluorine bonds (Zhang et al., 2021; Dombrowski et al., 2018).

Humans can be exposed to fluoride through several pathways, including drinking water, diet, consumer products, soil, and air. Approximately 40 to 70 percent of cumulative fluoride exposure in the United States is estimated to be via drinking water (U.S. EPA, 2010b). Many communities add fluoride (in the form of fluorosilicic acid and sodium hexafluorosilicate (ATSDR, 2003)) to the public drinking water supply to protect against dental caries (i.e., cavities). In geographic areas rich in naturally occurring fluoride-containing minerals, fluoride levels in groundwater, well water, and, rarely, surface water can exceed the target concentration for fluoride supplementation in community water systems (U.S. Department of Health and Human Services, 2015; Edmunds and Smedley, 2012). Fluoride has also been detected in dietary vegetation (e.g., dark, leafy greens and tea leaves), as plants can absorb fluoride from soil and fertilizers (Huang et al., 2025). Exposure of the general population to fluoride in air is expected to be low relative to exposure via ingestion, although exceptions could include populations living near industrial sites such as coal-burning facilities (U.S. EPA, 2010b; NRC, 2006). Exposure to fluoride can also occur through the use of dental care products (e.g., toothpaste, mouthwash) or fluoride-containing dietary supplements or medications (NIH, 2025; ATSDR, 2003).

2.1.2. Potential Human Health Effects of Fluoride Exposure

The focus of the EPA Fluoride Human Health Toxicity Assessment is the harmful human health effects of fluoride exposure and not the beneficial health effects. Fluoride exposure can be associated with both beneficial and harmful effects on human health, depending on the level of exposure. At lower exposure levels, fluoride decreases the prevalence of tooth decay (dental caries/cavities), which is one of the most common chronic diseases among American children (CDC, 2024). Higher exposure levels in children have been associated with multiple adverse effects including dental fluorosis, neurodevelopmental effects, and skeletal fluorosis (EFSA Scientific Committee, 2025; Taylor et al., 2025; Health Canada, 2024; NTP, 2024; Taher et al., 2024; Health Canada, 2010; U.S. EPA, 2010b; OEHHA, 2008; ATSDR, 2003; U.S. EPA, 1987). These effects are described briefly here as well as in Section 2.2.

Exposure to high levels of fluoride during childhood can affect the development of permanent (adult) teeth, resulting in a condition known as dental fluorosis. Dental fluorosis can range from very mild or mild (e.g., cosmetic impacts characterized by small white stripes or opaque areas) to severe (e.g., adverse effects on the tooth structure, including brown staining and pitting that is distinct from cavities or caries) (U.S. EPA, 2010a). Additionally, NTP's Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopment and Cognition (NTP, 2024) (*NTP monograph* herein) and the accompanying meta-analysis (Taylor et al., 2025) concluded that exposure to fluoride was associated with lower IQ in children, but that there was uncertainty in the relationship at levels of fluoride below 1.5 mg/L in drinking water. Prolonged exposure to high levels of fluoride is associated with an increased prevalence of bone weakening (skeletal fluorosis), which can include brittle bones, increased risk of fractures, and severe bone

abnormalities (defined as "crippling" or stage III skeletal fluorosis[1]) (U.S. EPA, 2010a; Dean, 1942).

2.1.3. SDWA National Primary Drinking Water Regulation for Fluoride

SDWA is the primary federal law protecting tap water provided to consumers by water systems across the country. SDWA was enacted in 1974 in response to "accumulating evidence that our drinking water contains unsafe levels of a large variety of contaminants." *Envtl. Def. Fund, Inc.* v. *Costle,* 578 F.2d 337, 339 (D.C. Cir. 1978). In passing SDWA, Congress intended to ensure "that water supply systems serving the public meet minimum national standards for protection of public health" (H.R. Rep. No. 93-1185, at 1 (1974)).

SDWA covers public drinking water sources including rivers, lakes, reservoirs, springs, and groundwater wells.[2] Under SDWA, the EPA is authorized to set national health-based standards for drinking water to protect against both naturally occurring and man-made contaminants that may be found in drinking water.

In the United States, fluoride in public water systems is regulated as a drinking water contaminant under SDWA. The EPA's current National Primary Drinking Water Regulation (NPDWR) for fluoride was established in 1986 (U.S. EPA, 1986b) with a maximum contaminant level (MCL) and maximum contaminant level goal (MCLG) of 4 mg/L to protect against crippling (stage III)[1] skeletal fluorosis. The EPA also established a secondary MCL of 2 mg/L to protect against cosmetically objectionable (i.e., moderate to severe) dental fluorosis in children. A secondary MCL is non-enforceable and may address aesthetic, technical, and/or cosmetic considerations.

Importantly, the EPA does not make recommendations about adding fluoride to drinking water. SDWA prohibits the EPA from requiring the addition of any substance to drinking water for preventive health care purposes (Section 1412(b)(11)) (U.S. EPA, 2019d). For communities that add fluoride to their public water systems, the U.S. Public Health Service recommends an optimal fluoride concentration of 0.7 mg/L to provide protection against dental caries while limiting the risks of harmful health effects, such as severe dental fluorosis (U.S. Department of Health and Human Services, 2015). Decisions about whether to add fluoride to drinking water are made on a state or local basis.

To date, the EPA has reviewed the current fluoride NPDWR four times as part of the Six-Year Review (SYR) of NPDWRs required under SDWA (Section 1412(b)(9)). Following SYR1 in 2003 (U.S. EPA, 2003), the EPA requested a review of fluoride health effects data by the National Research Council (NRC) of the National Academies of Science (NRC, 2006). Following its review, the NRC recommended that the EPA lower its MCLG of 4 mg/L for fluoride, concluding that the MCLG was not protective against severe dental fluorosis and may

[1] Stage III skeletal fluorosis is the most severe stage (often referred to as the "crippling" stage) and represents alterations in bone architecture and calcification of tissues that limit an individual's range of motion. Stage II skeletal fluorosis is the stage before mobility is significantly affected, and is characterized by chronic joint pain, arthritic symptoms, slight calcification of ligaments, and osteosclerosis of the cancellous bones (NRC, 2006).
[2] SDWA does not regulate private wells that serve fewer than 25 individuals.

not be protective against increased bone fractures. In response, in 2010, the EPA released an updated fluoride dose-response analysis for noncancer effects (focused on dental fluorosis and skeletal fluorosis, including skeletal fractures) and derived an RfD of 0.08 mg/kg/day (U.S. EPA, 2010b) based on the critical effect of severe dental fluorosis in a study in children (Dean, 1942). This RfD is protective against skeletal fluorosis (specifically, clinical stage II skeletal fluorosis) and skeletal fractures.

The EPA's 2010 fluoride dose-response analysis (U.S. EPA, 2010a) was not complete in time for SYR2 but it was used to derive a lower potential revised MCLG for SYR3. However, the NPDWR for fluoride was not deemed a candidate for revision at that time due to other NPDWRs of higher agency priority (U.S. EPA, 2017, 2016).

In SYR4, the EPA derived a potential revised MCLG of 0.9 mg/L but concluded that the NPDWR for fluoride was not a candidate for revision at that time due to ongoing higher priority actions and emerging research on neurodevelopmental effects of fluoride (U.S. EPA, 2023a). Specifically, the EPA noted that future regulatory decision-making may be informed by the NTP monograph (NTP, 2024), which had not been finalized at the time of SYR4 publication.

2.2. Critical Review of Recent Assessments on Fluoride Toxicity from National and International Agencies

As part of the assessment plan (U.S. EPA, 2026), to help inform the scope of the toxicity assessment, the EPA conducted a critical review of peer-reviewed assessments on fluoride published by multiple national and international agencies to review their findings about the types of harmful health effects that can result from fluoride exposure. The assessments conducted literature reviews, synthesized the available evidence, and developed conclusions about the potential harmful health effects of fluoride. They are briefly described below. They consistently reported associations between fluoride exposure and dental fluorosis and neurodevelopmental effects.

2.2.1. Noncancer Health Effects Conclusions of Recent Assessments on Fluoride

The noncancer health effects conclusions of the three most recent assessments on fluoride from national and international agencies are described briefly in this section. Additionally, a summary of noncancer toxicity values for chronic oral exposure to fluoride developed by the EPA and other federal or international health agencies is presented in Table 2-1.

NTP (2024) evaluated the evidence on fluoride exposure and nervous system effects (i.e., neurodevelopmental and cognitive effects). The NTP monograph included human, animal, and mechanistic (e.g., in vitro) studies published through May 1, 2020, and an addendum on findings from studies on children's IQ published through October 2023. The assessment applied rigorous and established systematic review methods (Office of Health Assessment and Translation methodology (NTP, 2019)) that included evaluation for risk of bias. NTP concluded with moderate confidence that exposure to fluoride was associated with lower IQ in children, but that there was uncertainty in the relationship at levels of fluoride below 1.5 mg/L in drinking water. The NTP monograph did not develop a toxicity value (e.g., RfD) for fluoride. A related,

subsequent meta-analysis of epidemiological data by NTP authors published in the peer-reviewed literature (Taylor et al., 2025) reported a quantitative estimate of the decrease in children's IQ per unit of fluoride exposure. The inverse association between fluoride exposure and IQ remained consistent when data were stratified by study quality (high or low risk of bias), sex, country, fluoride exposure matrix (measured in drinking water or urine), outcome assessment type, and prenatal or postnatal exposure but an RfD was not proposed.

The most recent and comprehensive final assessment of fluoride health effects was conducted by EFSA (EFSA Scientific Committee, 2025). The goal of the EFSA assessment was to determine whether there are health effects more sensitive than dental outcomes that would necessitate revisions to the European Union health-based guidance values for fluoride in food and drinking water. EFSA conducted a broad literature survey on the hazards of fluoride in humans and animals. From the large volume of literature that was identified, EFSA selected the following health outcome categories for full systematic review and hazard identification: neurotoxicity and developmental neurotoxicity, due to concerns raised in recent years regarding the inverse association between fluoride exposure and neurodevelopmental effects (e.g., NTP (2024)); bone health (e.g., bone mineral density; risk of bone fractures), which could be affected at lower exposure levels compared with skeletal fluorosis; and thyroid effects, which are of potential concern due to chemical similarities between fluoride and iodide (iodide is needed for the body to make thyroid hormones). EFSA did not perform hazard identification for dental or skeletal fluorosis because EFSA considered the association between these effects and fluoride exposure to be well-established in the scientific community. EFSA's assessment was performed using the EFSA systematic review framework and included critical appraisal of risk of bias (EFSA Scientific Committee, 2020).

Consistent with NTP's findings (NTP, 2024), EFSA found that total fluoride intake was associated with harmful health effects on the developing brain at fluoride drinking water concentrations greater than 1.5 mg/L. EFSA determined that the evidence was inconclusive for an association at lower exposure levels (EFSA Scientific Committee, 2025). EFSA concluded that dental fluorosis remained the most sensitive outcome in children up to eight years of age and derived a tolerable upper intake level (UL) of 1.4 mg/L for fluoride in drinking water, which EFSA anticipated to protect against all adverse effects including neurodevelopmental effects. The UL was based on the findings of Dean (1942), which is a study that describes fluoride exposure and dental fluorosis in children. Early childhood is considered the relevant exposure window for developing dental fluorosis, as it corresponds with enamel development in adult teeth. For children greater than eight years of age and adults including pregnant women, a reference point of 1.5 mg/L based on neurodevelopmental effects was used by EFSA to derive a safe intake level of fluoride. This level is also anticipated to protect against other potential adverse effects, such as effects on bone health, which EFSA determined are not anticipated to occur at fluoride drinking water concentrations less than 1.5 mg/L. With respect to potential thyroid toxicity, EFSA concluded that the available data provided insufficient evidence to support adverse effects of fluoride on thyroid function (EFSA Scientific Committee, 2025).

Health Canada also recently published a systematic review of the health hazards of fluoride exposure via drinking water (Taher et al., 2024). The results of the systematic review were

presented to an expert panel convened by Health Canada that would develop recommendations about dental fluorosis and neurocognitive development in children as sensitive endpoints of concern and develop a health-based value for fluoride in drinking water (Health Canada, 2024). Health Canada's review focused on updating conclusions from previous fluoride health assessments and used systematic review methods consistent with Cochrane guidelines (Higgins and Green, 2011) that included critical appraisal of risk of bias (these methods are also consistent with the EPA's systematic review methods). The panel agreed with the use of the Dean (1942) study for derivation of a health-based value for fluoride based on dental fluorosis. The panel also concluded that while neurocognitive effects are a sensitive endpoint of concern, there was an insufficient basis to recommend a health-based value for neurocognitive effects in children due to uncertainties regarding effects in the low-dose range (Health Canada, 2024).

Table 2-1. Summary of Available Noncancer Health Effects Toxicity Values for Chronic Oral Exposure to Fluoride Derived by the EPA and Other Agencies

Reference Value Source	Health Effect(s)	Reference Value(s)	Point of Departure	POD Type and Qualifier	Data Source	Total Uncertainty Factors
EFSA (EFSA Scientific Committee, 2025)	Mild to severe dental fluorosis	Infants (< 1 year): UL = 1 mg/day (0.14 mg/kg/day) Toddlers (1–3 years): UL = 1.6 mg/day (0.13 mg/kg/day) Children (4–8 years): UL = 2 mg/day (0.1 mg/kg/day)[c]	1.4 mg/L	BMDL (based on BMR = 5%)	Dean (1942)	1
	Central nervous system effects	"Safe intake level" for age groups > 8 years old and adults, including pregnant women = 3.3 mg/day[d]	1.5 mg/L	Reference point above which adverse effects are consistently observed	Based on weight of evidence analysis	N/A
EPA OW (U.S. EPA, 2010b)	Severe dental fluorosis	RfD = 0.08 mg/kg/day[a]	1.87 mg/L	BMDL (based on BMR = 0.5%)	Dean (1942)	1
Health Canada (Health Canada, 2010)	Moderate to severe dental fluorosis	TDI = 0.105 mg/kg/day[e]	1.56 mg/L	BMDL (based on BMR = 1%)	Dean (1942)	1
California OEHHA (OEHHA, 2008)	Moderate to severe dental fluorosis	REL = 0.04 mg/kg/day[g]	1 mg/L	NOAEL	Dean (1942)	1
ATSDR (ATSDR, 2003)	Skeletal effects (increased risk of bone fractures)	MRL = 0.05 mg/kg/day[f]	0.15 mg/kg/day	NOAEL	Li et al. (2001)	3 (to account for human variability)
EPA IRIS (U.S. EPA, 1987)	Moderate and severe dental fluorosis	RfD = 0.06 mg/kg/day[b]	1 mg/L	NOAEL	Hodge (1950) cited in Underwood (1977)	1

Reference Value Source	Health Effect(s)	Reference Value(s)	Point of Departure	POD Type and Qualifier	Data Source	Total Uncertainty Factors
EPA OW National Primary Drinking Water Regulation (U.S. EPA, 1986b)	Severe skeletal fluorosis (stage III)	MCL and MCLG = 4 mg/L[h]	10 mg/L	NOAEL	(Koop, 1984; WHO, 1984; Shapiro, 1983)	2.5 (safety factor)

EPA = Environmental Protection Agency; POD = point of departure; EFSA = European Food Safety Authority; UL = tolerable upper intake level; BMDL = benchmark dose lower confidence limit; BMR = benchmark response; N/A = not applicable; OW = Office of Water; RfD = reference dose; TDI = tolerable daily intake; OEHHA = Office of Environmental Health Hazard Assessment; REL = reference exposure level; NOAEL = no-observed-adverse-effect level; ATSDR = Agency for Toxic Substances and Disease Registry; MRL = minimal risk level; IRIS = Integrated Risk Information System.

[a]The EPA OW RfD was calculated by summing the estimated fluoride intake by children from drinking water at the BMDL (0.07 mg/kg/day, based on drinking water intake rates and body weights from the United States Department of Agriculture 1977/1978 Nationwide Food Consumption Survey (Ershow and Cantor, 1989) and estimated dietary fluoride intake of 0.01 mg/kg/day (from McClure (1943)).

[b]The EPA IRIS RfD was calculated by summing the estimated fluoride intake by children from drinking water at the NOAEL (0.05 mg/kg/day, based on drinking water intake of 1 L/day for a 20-kg child) and an estimated dietary fluoride intake of 0.01 mg/kg/day (from McClure (1943)).

[c]The EFSA ULs were calculated by summing the estimated fluoride intake by children from drinking water at the BMDL (1.12, 1.46, and 1.79 mg/day for children ages 6–12 months, 1–3 years, and 4–8 years; based on drinking water intakes of 0.8–1, 1.3, and 1.6 L/day for the three age groups and assuming that 20% of fluids will contain negligible fluoride levels) and 1940s era estimates of dietary fluoride intake (0.09, 0.12, and 0.20 mg/day for children ages 6–12 months, 1–3 years, and 4–8 years), with default body weights of 8.8, 12, and 20 kg for age groups 6–12 months, 1–3 years and 4–8 years, respectively. This is documented in Table 27 of EFSA Scientific Committee (2025).

[d]The EFSA "safe intake level" uses 1.5 mg/L in drinking water as the reference point above which there is consistent evidence of adverse neurodevelopmental effects in children. The safe intake level was calculated by summing the estimated fluoride intake from drinking water at 1.5 mg/L (2.4 mg/day, based on drinking water intake rate of 2 L/day for adult women and assuming that 20% of fluids contain negligible fluoride levels) with estimated fluoride intake from food (0.64 mg/day) and dental products (0.3 mg/day).

[e]The Health Canada TDI was calculated by summing the estimated fluoride intake by children from drinking water at the BMDL (98.5 µg/kg/day, based on 0.8 L/day for a 13-kg child) with estimated fluoride intake from food (5.4 µg/kg/day assuming a 1940s diet for a 1-to-4-year-old child), soil (1.19 µg/kg/day), and air (0.01 µg/kg/day).

[f]The ATSDR MRL was calculated based on the total daily fluoride intake reported by Li et al. (2001) (3 mg/day at the NOAEL) and a reference body weight of 55 kg. An uncertainty factor of 3 was applied to account for variability in the human population (decreased from 10 because the NOAEL was derived from a sensitive subpopulation (elderly men and women)).

[g]The California OEHHA REL was calculated from estimated fluoride intake by children from drinking water at the NOAEL, based on the assumption that an 18-kg child drinks 720 mL of water per day.

[h]An RfD was not derived. Instead, a safety factor equivalent to 2.5 was applied to the NOAEL of 10 mg/L based on crippling skeletal fluorosis to calculate the MCLG of 4 mg/L. See 50 FR 47142.

2.2.2. Cancer Conclusions of Assessments on Fluoride

The EPA also reviewed assessments published by multiple national and international agencies that evaluated the potential carcinogenicity of fluoride (Table 2-2). Among the published assessments, there are consistent conclusions that the available data from epidemiology and animal studies are insufficient to support a cancer determination for fluoride. NRC (2006) reviewed the available information on carcinogenicity and fluoride exposure at the request of the EPA. The NRC (2006) determined that the evidence was "tentative and mixed" and recommended that additional highly focused epidemiological studies of fluoride and cancer be conducted to assess specific tissues for which there were suggestions of carcinogenic activity in the available the human and animal literature (i.e., osteosarcomas and cancers of the buccal cavity, kidney, bones, and joints). More recently, EFSA concluded that the available evidence did not support an association between fluoride exposure and bone cancer (EFSA Scientific Committee, 2025), and the systematic review conducted by Health Canada did not identify an association between fluoride exposure and cancer (Taher et al., 2024). In summary, published assessments of the potential carcinogenicity of fluoride conclude that the evidence from human and animal studies is insufficient or unclassifiable with respect to cancer. Therefore, the EPA will not consider cancer in the future Fluoride Human Health Toxicity Assessment.

Table 2-2. Cancer Conclusions from Exposure to Fluoride by the EPA and Other Agencies

Source	Overview/Scope	Basis for Cancer Conclusion	Cancer Conclusions
European Food Safety Authority (EFSA) (EFSA Scientific Committee, 2025)	Toxicity assessment of fluoride in drinking water and food, mandated by the European Commission to update the previous EFSA fluoride assessment from 2005. Systematic review performed for bone cancer but not other types of cancer.	Descriptive conclusion based on the weight of evidence. No specific approach for deriving a cancer descriptor was followed.	Lack of association between fluoride exposure and bone cancer. Other types of cancer not addressed.
California Office of Environmental Health Hazard Assessment (OEHHA) (OEHHA, 2008)	Assessment of the evidence of carcinogenicity to inform deliberations on whether fluoride and its salts should be listed under California Proposition 65. Conclusions based on studies included in previous reviews of fluoride carcinogenicity by NRC (1993), NRC (2006), and ATSDR (2003) and supplemented by a relevant human epidemiological study published after NRC 2006 (Bassin et al., 2006).	Descriptive conclusion based on the weight of evidence. No specific approach for deriving a cancer descriptor was followed.	The California Proposition 65 Carcinogen Identification Committee concluded "fluoride and its salts has not been clearly shown to cause cancer" (OEHHA, 2011).
Health Canada[a] (Health Canada, 2010)	Technical guidance document for fluoride used to develop Guidelines for Canadian Drinking Water Quality.	Health Canada criteria for classification of carcinogenicity.[b] Group VI corresponds to IARC Group 3, which is used when evidence of carcinogenicity is inadequate in humans and inadequate or limited in experimental animals.	*Group VI: Unclassifiable with respect to carcinogenicity in humans*.
EPA Office of Pesticide Programs (OPP) (U.S. EPA, 2007, 1996b)	Reregistration Eligibility Decisions for the manufacture and use of sodium fluoride as a fungicide (used as a wood preservative in lumber and other wood products), and for the manufacture and use of cryolite (sodium aluminofluoride) as an insecticide on food crops.	Classification system described in the EPA's 1986 Guidelines for Carcinogen Risk Assessment (U.S. EPA, 1986a).[c] Group D is generally used for agents with inadequate human and animal evidence of carcinogenicity or for which no data are available.	Sodium fluoride and sodium aluminofluoride are described as *Group D (not classifiable as to human carcinogenicity).*
National Research Council (NRC) (NRC, 2006)	Review of fluoride data by the NRC of the National Academies of Science based on the request from the EPA OW to assess the adequacy of the current drinking water standards under SDWA to protect human health.	Descriptive conclusion based on the weight of evidence. No specific approach for deriving a cancer descriptor was followed.	Based on "consideration of data from humans, genotoxicity assays, and studies of mechanisms of action in cell systems (e.g., bone cells in vitro), the evidence on the potential of fluoride to initiate or promote

Source	Overview/Scope	Basis for Cancer Conclusion	Cancer Conclusions
			cancers, particularly of the bone, is tentative and mixed."
World Health Organization (WHO) (WHO, 2004)	Background document for the development of guidelines for drinking water quality for fluoride by WHO.	The document does not provide an independent conclusion but instead only notes agreement with IARC and IPCS.	The document notes agreement with the IARC conclusion of inadequate evidence of carcinogenicity in humans (IARC, 1987b) and the IPCS conclusion "that the evidence of carcinogenicity in laboratory animals is inconclusive and that the weight of evidence does not support the hypothesis that fluoride causes cancer in humans; however, the data on bone cancer are relatively limited" (IARC, 1987a).
Agency for Toxic Substances and Disease Registry (ATSDR) (ATSDR, 2003)	Toxicological profile assessment for fluorides, hydrogen fluoride, and fluorine, including an assessment of cancer.	Neither a descriptive conclusion nor a formal cancer descriptor was provided.	Did not provide a conclusion but noted data gaps. "Additional studies are needed to further evaluate the potential of fluoride to induce bone cancers following chronic oral exposure."
International Agency for Research on Cancer (IARC) (IARC, 1987b)	Determination of cancer hazard for fluorides (inorganic, used in drinking water).	IARC Monograph hazard classifications.[d] Group 3 is used for agents, mixtures, and exposure circumstances for which the evidence of carcinogenicity is inadequate in humans and inadequate or limited in experimental animals.	Fluorides (inorganic, used in drinking water) are described as *Group 3: Not classifiable as to its carcinogenicity to humans*.

EPA = Environmental Protection Agency; EFSA = European Food Safety Authority; OEHHA = Office of Environmental Health Hazard Assessment; NRC = National Research Council; ATSDR = Agency for Toxic Substances and Disease Registry; IARC = International Agency for Research on Cancer; OPP = Office of Pesticide Programs; OW = Office of Water; SDWA = Safe Drinking Water Act; WHO = World Health Organization; IPCS = International Programme on Chemical Safety.

[a]Health Canada recently conducted a systematic review (Taher et al., 2024); they did not find additional evidence of an association between fluoride exposure and cancer, based on studies published between 2016 and 2021.

[b]Health Canada categories for classification of carcinogenicity are as follows: Group I (carcinogenic to humans), Group II (probably carcinogenic to humans), Group III (possibly carcinogenic to humans), Group IV (unlikely to be carcinogenic to humans), Group V (probably not carcinogenic to humans), and Group VI (unclassifiable with respect to carcinogenicity in humans).

[c]U.S. EPA (1986a) categories for classification of carcinogenicity are as follows: Group A (carcinogenic to humans), Group B (probably carcinogenic to humans), Group C (possibly carcinogenic to humans), Group D (not classifiable as to human carcinogenicity), and Group E (evidence of noncarcinogenicity for humans).

[d]IARC monographs' hazard classifications for carcinogenicity are as follows: Group 1 (carcinogenic to humans), Group 2A (probably carcinogenic to humans), Group 2B (possibly carcinogenic to humans), Group 3 (not classifiable as to its carcinogenicity to humans), and Group 4 (probably not carcinogenic to humans).

2.3. Scoping and Problem Formulation Summary

The objective of the toxicity assessment is to derive a noncancer oral RfD for fluoride to protect human health (Section 1). Figure 2-1 shows the Fluoride Human Health Toxicity Assessment development process and Figure 2-2 summarizes the EPA's critical review of other national and international agencies' assessments on fluoride that were used to inform the scope of the toxicity assessment.

The scope of the EPA's fit-for-purpose toxicity assessment will be dental fluorosis and neurodevelopmental effects. This decision is based on the EPA's critical review of peer-reviewed assessments published by national and international agencies that identified consistent findings about these two harmful health effects in children (Section 2.2.1). These assessments used well-established systematic review methods consistent with Gold Standard Science (United States Executive Office of the President Donald Trump, 2025) and the EPA's approach to systematic review, including risk-of-bias evaluations of the health effects studies informing hazard determination. Therefore, the EPA is accepting the consistent hazard conclusions related to dental fluorosis and neurodevelopmental effects, which have been identified by multiple agencies as well-established and sensitive adverse effects associated with fluoride exposure. Leveraging well-established hazard conclusions allows the EPA to expedite the toxicity assessment development process by focusing resources on the dose-response analysis. The dose-response analysis will be conducted independently, based on the methodology described in this protocol. The health benefits of oral fluoride exposure (e.g., decreased dental caries) are not within the scope of the toxicity assessment.

The toxicity assessment will not include an evaluation of fluoride exposure and cancer. The EPA's review of published health assessments on fluoride and cancer concluded that the evidence to support carcinogenicity of fluoride was either weak or insufficient to make a determination (Section 2.2.2). In addition, the results of the EPA's preliminary literature survey (U.S. EPA, 2026), which was conducted as part of the assessment plan, indicate that newly identified literature on fluoride and cancer is not anticipated to impact existing conclusions about cancer. Taken together, the EPA's critical review of existing assessments and the EPA's preliminary literature survey results showed that a reassessment of carcinogenicity would be unlikely to yield a conclusion different from those of the existing published assessments.

Since the EPA is accepting well-established hazard conclusions about dental fluorosis and neurodevelopmental effects from fluoride exposure, hazard identification systematic review steps (i.e., evidence synthesis and integration) will not be performed for the toxicity assessment. Additionally, the acceptance of hazard conclusions does not imply that the adverse effects can be observed at all dose/exposure levels; a dose-response analysis is needed to characterize the dose or exposure levels at which the hazards occur. Thus, as outlined in this protocol, the EPA is focusing its resources and expertise on systematic literature searching and screening to identify the most suitable studies for dose-response analysis for these two adverse children's health outcomes (Sections 4 and 5). The EPA will perform study evaluation and data extraction for the studies identified as most suitable for dose-response analysis (Sections 6 and 7) and will perform dose-response analyses on quality studies for toxicity value derivation (Section 8).

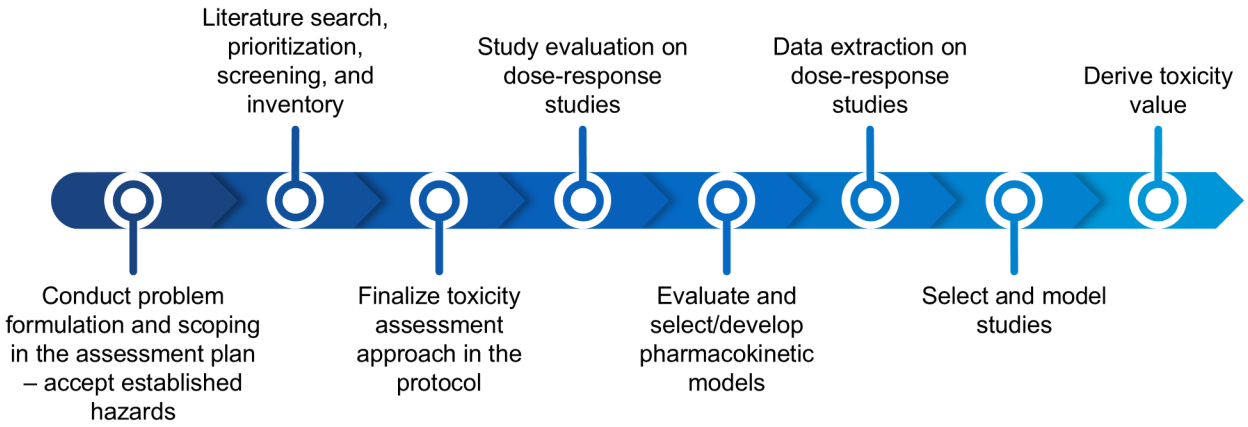


Figure 2-1. Stages of the EPA's Fluoride Human Health Toxicity Assessment Development Process

The toxicity assessment will focus on human (epidemiological) studies only. Fluoride's large epidemiological health effects database allows the toxicity assessment to focus on human studies. Furthermore, fluoride toxicity values derived in final assessments published by the EPA and other agencies have all been based on epidemiological data (Table 2-1), which provides additional support for scoping to human studies. Additionally, the use of epidemiological studies for toxicity value derivation eliminates uncertainties associated with extrapolation from animals to humans and for this reason, human data are preferred over animal data when robust human studies are available (U.S. EPA, 2022b).

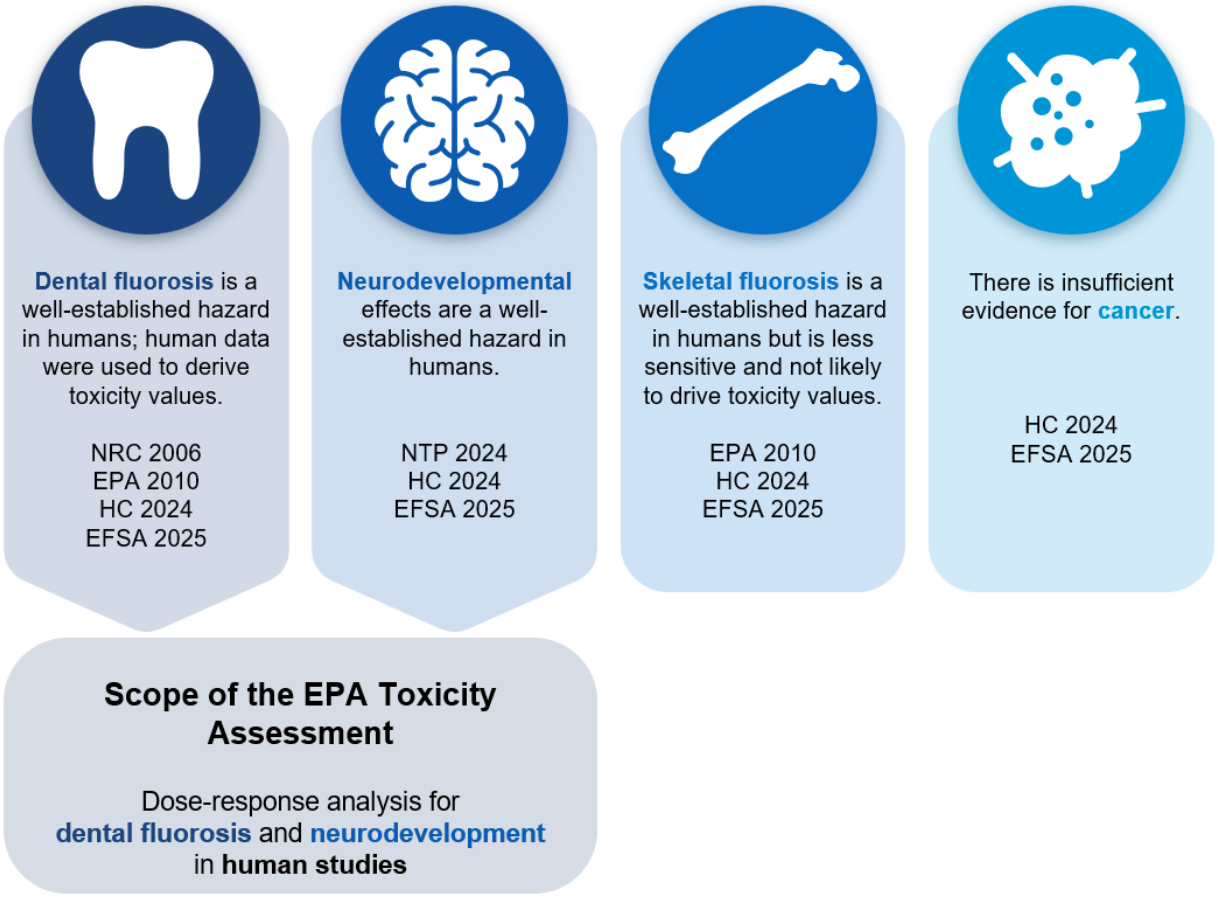


Figure 2-2. Summary of the Scope of the EPA's Future Fluoride Human Health Toxicity Assessment

NRC = National Research Council; EPA = Environmental Protection Agency; HC = Health Canada; EFSA = European Food Safety Authority; NTP = National Toxicology Program.

2.4. Key Science Issues

Key science issues describe areas of scientific complexity and uncertainty for fluoride that were identified in the assessment plan. The EPA has considered these key science issues and here in the protocol, provides specifics about how these issues will be considered and addressed in the toxicity assessment. The EPA identified four key scientific issues based on its critical review of recent fluoride assessments (Section 2.2.1) and the results of the preliminary literature survey (Sections 4.3 and 4.4; (U.S. EPA, 2026)):

Validity of exposure metrics in epidemiology studies.

Science Issue: Exposure to fluoride can be estimated in many ways, including measuring sources of fluoride that are outside of the body (e.g., concentrations of fluoride in drinking water) or measuring the levels of fluoride inside the body (e.g., concentrations of fluoride in urine, blood, or other tissues, which are known as biomarkers of fluoride exposure).

Addressing the Science Issue: For the toxicity assessment, fluoride concentrations in drinking water as well as internal biomarker measures of fluoride (e.g., urine concentrations) are

considered valid approaches for characterizing human fluoride exposure, but the strengths and limitations of the exposure measures in each study are evaluated. Specifically, during study evaluation, the analytical methods used for all exposure measures are considered. For example, fluoride in drinking water or urine can be measured with multiple laboratory techniques with varying sensitivity, specificity, and susceptibility to interferences. Analytical measurement techniques used and subsequent potential for exposure measurement error are considered. Additionally, the timing of exposure measurement should coincide or overlap with the critical window of exposure for the health outcome of interest. Also, measurements of fluoride in urine or other matrices are subject to within- and between-person variability due to factors including changing exposure over a lifetime, and inter-individual differences in fluoride elimination from the body. If needed, additional analyses using relevant information from the health effects and absorption, distribution, metabolism, and excretion (ADME) literature surveys are conducted to inform these study evaluation considerations. Thus, study evaluation will also consider whether the reported exposure biomarkers can be used to estimate oral fluoride intake during an etiologically relevant time period, based on knowledge of fluoride pharmacokinetics and lifestage sensitivities. Criteria for the consideration of exposure metrics are described in Section 6.2.1.

Completeness of exposure characterization in epidemiology studies.

Science Issue: Related to the validity of different exposure metrics, relying on studies with only measures of fluoride concentrations in drinking water can be a challenge. Drinking water fluoride concentration data only provide an estimate of a single exposure source and often require assumptions that can be a source of measurement error. And, since oral exposure to fluoride can occur from other sources besides drinking water, additional steps are necessary to relate drinking water fluoride exposure to total fluoride exposure. The EPA notes that the date of a publication needs to be considered when evaluating fluoride exposure. For instance, the Dean (1942) study that was previously used to derive toxicity reference values for fluoride by the EPA (U.S. EPA, 2010a) (Table 2-1) used fluoride concentration in drinking water as the exposure measure, and was conducted at a time before fluoride was added to dental products, so oral fluoride exposure in children in this study would have occurred exclusively through drinking water and (to a lesser extent) food intake. More recent studies of dental fluorosis and neurodevelopmental effects of fluoride were conducted in time periods during which significant fluoride exposure from dental products could contribute to total fluoride intake.

Addressing the Science Issue: For post-1940s epidemiology studies that evaluate outcomes only in relation to fluoride concentrations in drinking water, the estimation of total fluoride intake will need to account for other sources of fluoride exposure (e.g., dental products such as toothpaste).

Measurements of fluoride in urine (Rango et al., 2017), serum (Sankhala et al., 2014), or tissues, such as bone (Rugg-Gunn et al., 2011), teeth (Vieira et al., 2005), hair (Parimi et al., 2013), or nail clippings (Elekdag-Turk et al., 2019; Rango et al., 2017) are biomarkers of a person's total fluoride intake (Lavalle-Carrasco et al., 2021) meaning they account for all possible sources of fluoride exposure (e.g., drinking water, diet, dental products). However, some of these biomarkers (e.g., fluoride concentration in blood or urine) could be highly influenced by recent

exposure and might not represent an individual's typical or average exposure. Bone fluoride measurement is highly invasive and rarely performed in children and pregnant women, which are considered susceptible populations. Fluoride measurement in teeth might be highly influenced by topical exposure to dental products and might not reflect systemic fluoride or total fluoride exposure. Additionally, nail clippings represent average systemic concentrations over a long period of time and inter-individual differences in the uptake of fluoride to fingernails and toenails can substantially influence these measurements. In the toxicity assessment, study evaluations are customized to consider potential methodological concerns associated with the measurement of fluoride biomarkers (e.g., fluoride concentration in blood or urine) or measurements of fluoride in food or water. After study evaluation, the degree to which the exposure measure provides a complete exposure characterization is considered in the selection of studies for dose-response analysis. In addition, these different exposure measures will be evaluated for their ability to support estimation of oral fluoride intake relevant for RfD derivation.

Understanding how fluoride moves through the body.

Science Issue: Pharmacokinetics (PK) is used to describe how fluoride enters, moves through, and is eliminated from the body (Tan et al., 2020). This involves conducting a comprehensive review of the ADME data, as well as assessing mathematical models that represent these fluoride ADME processes. Key aspects of fluoride ADME data relevant to the future toxicity assessment include the age-dependent preferential distribution of fluoride ion to bone (Richards et al., 1994), the absence of further biotransformation in the body through metabolism, and the primary route of excretion via urine (Institute of Medicine, 1997b).

PK models are a mathematical representation of ADME processes. They are routinely used in EPA toxicity assessments (U.S. EPA, 2024a, b) and represent the preferred approach for relating external exposure with internal measurements. The use of a PK model allows for RfD derivation based on fluoride concentrations that are measured within the body (e.g., in blood or urine) (U.S. EPA, 2014b, 2006).

Currently, the EPA has identified two PK models (Jean et al., 2018; Rao et al., 1995) that have been developed to predict internal fluoride dosimetry while accounting for bone uptake and renal clearance.

Addressing the Science Issue: The EPA is developing a robust PK analysis to better understand how fluoride moves through the body and address the key science issues related to the validity of exposure metrics and exposure characterization. The EPA is considering existing PK models and approaches identified through its systematic review of the literature or public comment. This includes the approach adopted by EFSA (EFSA Scientific Committee, 2025), which used empirical correlations between total daily fluoride intake and fluoride urinary measurements. PK approaches are evaluated based on the fit to in vivo human data, the accuracy of the biological and mechanistic bases for the model, and the suitability of the model for addressing the toxicity assessment context for fluoride (specifically, whether lifestage-specific features of ADME are accurately described).

Adversity of dental fluorosis.

Science Issue: Dental fluorosis has been widely studied and has been used as the basis for fluoride toxicity values in final assessments developed by national and international agencies (Table 2-1), but there are differences in how adversity is defined across the range of severity categories for dental fluorosis. The severity levels of dental fluorosis are generally categorized, according to the Dean's Index, as *very mild* or *mild*, *moderate*, or *severe* (NRC, 2006), and these categories of increasing severity were used across fluoride assessments published by national and international agencies (Section 2.2). However, there are differences in the assignment of adversity to the different severity levels. The assessments that were critically reviewed by the EPA all agreed that *severe* dental fluorosis (characterized by dark yellow to brown staining and pitting) is an adverse health effect because enamel is lost in affected teeth. However, the assessments did not have similar conclusions about the adversity of *moderate* fluorosis. The EPA's 2010 dose-response analysis for fluoride (U.S. EPA, 2010a) was therefore based on the occurrence of *severe* dental fluorosis and considered *mild* and *moderate* severity to be primarily an aesthetic concern and not an adverse health effect. In contrast, the more recent EFSA (EFSA Scientific Committee, 2025) and Health Canada (Taher et al., 2024) assessments considered *moderate* dental fluorosis to be adverse based on potential psychological effects resulting from aesthetic effects of dental fluorosis. The EFSA dose-response analysis (EFSA Scientific Committee, 2025) was based on the combined incidence of *mild*, *moderate*, and *severe* dental fluorosis.

Addressing the Science Issue: The EPA is considering the adversity of different severities of dental fluorosis, primarily *moderate* and *severe*, based on the health effects literature, U.S. policy and precedent, and dose-response analysis approaches. The rationale for dose-response analysis decisions, including selection of appropriate benchmark response (BMR) values and identification of an RfD for fluoride, will be documented in the toxicity assessment.

3. Toxicity Assessment Objective and Specific Aims

Section Overview: This section of the protocol outlines the overall objective and specific aims of the Fluoride Human Health Toxicity Assessment. The overall objective is to identify the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects. This amount is referred to as a toxicity value or RfD. To derive a toxicity value, the EPA will follow specific steps (i.e., specific aims) to identify and evaluate studies that assess associations between fluoride exposure and dental fluorosis or neurodevelopmental effects in humans. The EPA will then use the best-available science to develop a toxicity value for fluoride.

The objective of the toxicity assessment is to develop an oral RfD for fluoride. The EPA will accomplish this objective through the following specific aims:

- **Literature search.** Conduct a comprehensive literature search to identify epidemiological studies that evaluate fluoride exposure and neurodevelopmental effects or dental fluorosis.
- **Literature screening.** Review the literature search results against inclusion criteria (known as populations, exposures, comparators, and outcomes or PECO criteria; described in Table 4-1 and Section 4) to identify studies likely to be relevant to the toxicity assessment.
- **Preliminary data extraction and health effects literature inventory development.** Extract study design information from studies meeting the inclusion criteria to develop literature inventories that characterize the extent of the evidence (known as systematic evidence maps).
- **Preliminary PK/PBPK literature inventory development.** Identify studies reporting PK information or PBPK models for fluoride in humans, which may be used to inform dose-response analysis.
- **Study evaluation.** Conduct study evaluation of epidemiological studies that meet the PECO criteria (Table 4-1) and that provide quantitative information that is likely to support dose-response analysis (such as regression coefficients presented with statistical measures of variation). Study evaluation will consider factors that might impact the strength or direction of the observed effects (i.e., potential risk of bias) and factors that might limit the ability of a study to detect an association when one exists (i.e., sensitivity).
- **PK/PBPK model evaluation.** Conduct scientific and technical review of PK/PBPK models for use in the fluoride pharmacokinetic analysis. A fluoride PK/PBPK model (either sourced from the existing literature or developed by the EPA) will be considered for integration with the dose-response analysis to aid in estimating total fluoride exposure, especially when only limited external and biological measures of fluoride are available (e.g., fluoride concentration in drinking water and urine).
- **Dose-response analysis study selection and data extraction.** Select relevant epidemiological studies for dose-response analysis based on study evaluation conclusions

and consideration of dose-response analysis suitability criteria. Extract data on relevant health outcomes from the selected studies.

- **RfD derivation.** Derive an RfD from the available robust epidemiological data using best available dose-response analysis approaches. This includes the application of uncertainty factors to account for uncertainty and/or variability (e.g., differences among people and exposure durations as well as gaps in available data).
- **Characterize uncertainties and limitations.** Characterize uncertainties and identify key data gaps and research needs, such as limitations of the evidence, limitations of the systematic review, and consideration of dose relevance and PK information.

The toxicity assessment will use systematic review methods to transparently identify and evaluate relevant health effects literature for fluoride. An overview of the process is provided in Figure 3-1.

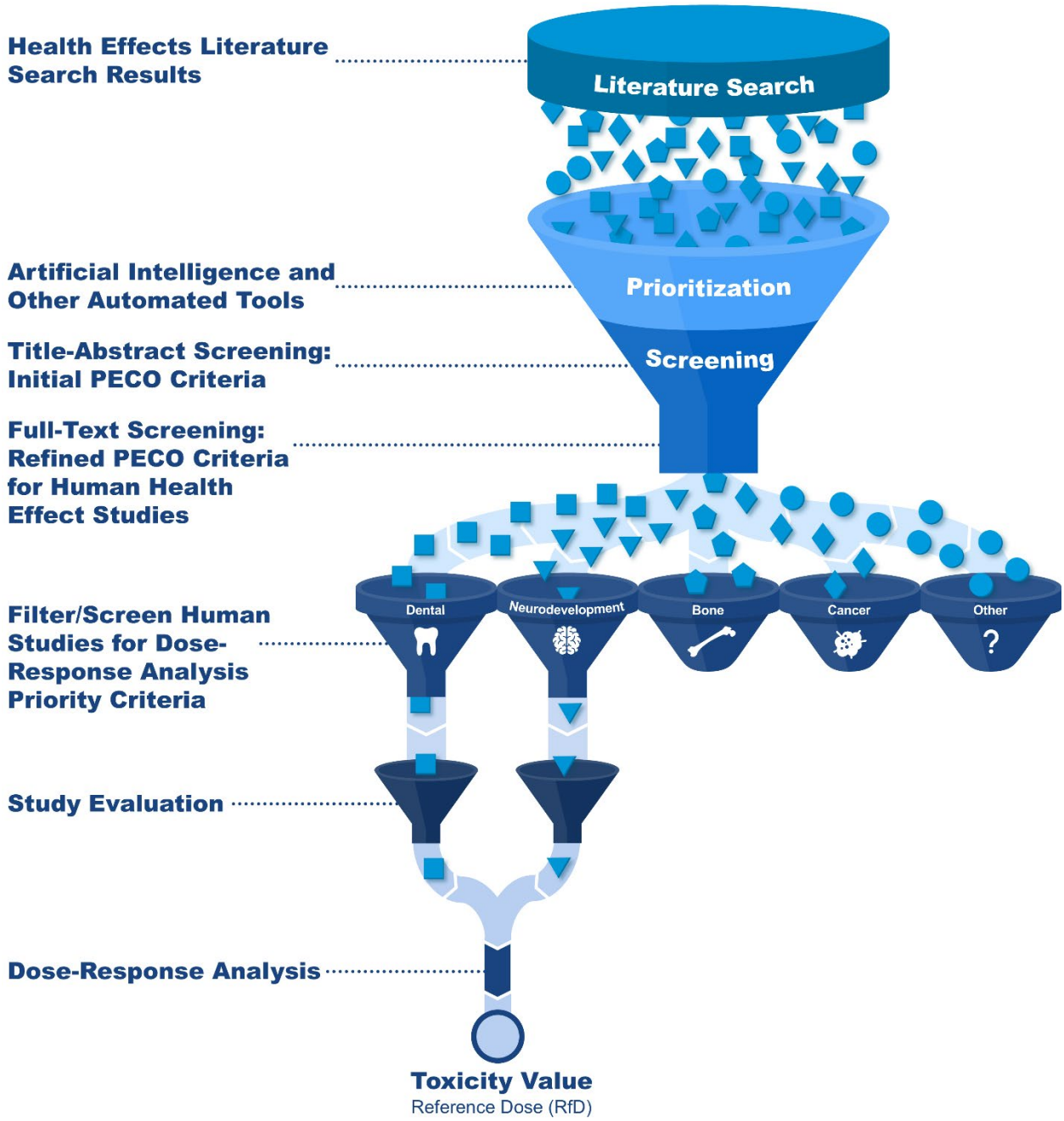


Figure 3-1. Overview of Fluoride Human Health Toxicity Assessment Literature Selection Process

PECO = populations, exposures, comparators, and outcomes; RfD = reference dose.

4. Literature Search and Screening

Section Overview: This section describes how the EPA finds (Section 4.1) and chooses (Section 4.2) the most relevant and best-available studies for the toxicity assessment of fluoride. This section also provides preliminary results for human health effects studies (Section 4.3) and studies describing how fluoride moves through the human body (i.e., pharmacokinetic or PK studies) (Section 4.4) from the initial literature searches that the EPA performed. The initial searches identified 268,093 studies. These studies were screened and sorted using a combination of automated tools (including artificial intelligence) and manual screening and checks by the EPA. The EPA identified 671 relevant human health effects studies, along with some PK studies, that will be considered further for the toxicity assessment.

As described in the assessment plan (U.S. EPA, 2026), the EPA performed an initial literature search of available peer-reviewed scientific literature databases and other resources (i.e., crosswalks with existing assessments and studies submitted through public comment). The published literature on fluoride is vast, encompassing nearly a century of data in humans, animals, and in vitro models, as well as pharmacokinetic data to inform the transport and elimination of fluoride in the body. The EPA conducted prioritization, screening, and sorting of the initial literature search results to identify studies most likely to be informative to the toxicity assessment and subsequent RfD derivation. Specifically, the EPA reviewed health effects studies to identify those with data relevant to the toxicity assessment (described in Section 4.2) and also amenable to dose-response analysis (Section 8.1), and screened ADME studies to identify those with relevant PK data that could potentially inform key science issues and RfD derivation (Appendix C.4). An updated literature search and screening is performed during toxicity assessment development, according to the methods outlined in the following sections.

4.1. Literature Search Strategies

Section Overview: The EPA identifies studies that may be useful for the toxicity assessment using systematic review methods that clearly and transparently document how studies are found and sorted. Initial broad searches of scientific literature databases were performed; an updated search is performed during toxicity assessment development (Section 4.1.1). The EPA also reviewed lists of studies cited in fluoride reports published by the EPA and other agencies against initial literature database search results ("crosswalking") to ensure that all relevant studies from these reports were captured (Section 4.1.2).

4.1.1. Database Searches

Database searches are conducted in PubMed and Web of Science databases using the search terms listed in Appendix C.1. The search terms are designed to be inclusive of all English language studies that mention fluoride or its synonyms/variations. The initial database searches

were conducted in November 1, 2024, as part of the assessment plan without publication date limitations, and the results were filtered to prioritize English language studies. A database search update is performed during toxicity assessment development to identify the most recently published literature and maintain an up-to-date literature inventory. The updated literature search is expected to be conducted within months of the planned release of the draft toxicity assessment for public comment.

4.1.2. Crosswalks with Existing Assessments and Other Sources

The EPA conducted crosswalks of the initial database search results with published fluoride assessments and other high-quality published references. This process (outlined in the following subsections) ensured that relevant literature and key studies identified in published fluoride assessments were included for review.

4.1.2.1. Health Effects Studies Cited in Existing Assessments and Other Sources

Lists of health effects studies cited in published fluoride assessments or systematic reviews by the EPA or other national and international agencies were compared against the initial database search results to ensure that these studies had been captured in the initial literature search results. The EPA conducted crosswalks of references cited in the following assessments (which were selected either because they were used to inform the EPA's 2010 dose-response assessment of fluoride (U.S. EPA, 2010b) or published after the 2010 assessment):

- Fluoride Dose-Response Analysis for Non-Cancer Effects – Dental Fluorosis: Evaluations of Key Studies (U.S. EPA, 2008b)
- Fluoride Dose-Response Analysis for Non-Cancer Effects – Fluoride-Related Skeletal Effects: Evaluations of Key Studies (U.S. EPA, 2008c)
- Fluoride Dose-Response Analysis for Non-Cancer Effects (U.S. EPA, 2010b)
- Fluoride in Drinking Water: A Scientific Review of EPA's Standards (NRC, 2006)
- NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopment and Cognition: A Systematic Review (NTP, 2024)
- Guidelines for Canadian Drinking Water Quality: Guideline Technical Document – Fluoride (Health Canada, 2010)
- Systematic review of epidemiological and toxicological evidence on health effects of fluoride in drinking water (Taher et al., 2024) associated with the 2023 Health Canada Panel Report (Health Canada, 2024)
- Updated consumer risk assessment of fluoride in food and drinking water including the contribution from other sources of oral exposure (EFSA Scientific Committee, 2025)

Additionally, fluoride literature identified through the EPA's SYR4 (U.S. EPA, 2023a) process as well as studies referenced in the American Dental Association Letter to the EPA Administrator (American Dental Association, 2025) were reviewed.

Studies identified in the above-listed publications that had not been identified in the initial database searches were added to the literature review and advanced directly to title-abstract screening.

4.1.2.2. Studies Submitted to the EPA During Public Comment Periods

Peer-reviewed references submitted during the EPA's public comment period for the assessment plan or the draft toxicity assessment are reviewed for relevance to the toxicity assessment and any not captured by the literature searches are added to the literature review. With regard to references provided to the EPA via the public comment period for the assessment plan, those published after the initial literature search cutoff date (November 1, 2024) are screened as part of a planned updated literature search for the toxicity assessment. Any references submitted during the assessment plan public comment period that had also been captured in the initial literature search but then excluded during an automated prioritization step (Section 4.2.3) are manually reviewed for relevance.

4.2. Literature Screening Strategies for Health Effects Studies

Section Overview: This section describes the methods and tools that the EPA uses to review the literature search results to identify studies that will be used in the assessment. The initial database searches performed by the EPA were broad and designed to find all studies with potentially useful information. As a result, 268,093 studies were identified in the initial database searches. To efficiently identify relevant studies from among these hundreds of thousands of studies, the EPA used automated approaches, including machine learning-based artificial intelligence (AI) and other tools (Section 4.2.3 and Appendix C.2) (known as "prioritization"). These approaches are established systematic review methods that have been validated and have been used for other assessments developed by the EPA. The results of the prioritization approaches were reviewed by experts.

Following the initial round of prioritization, scientists manually reviewed the title and abstract of each prioritized study against a set of criteria (Table 4-1) to confirm whether the study contains relevant health effects information (Sections 4.2.1 and 4.2.4) or other potentially relevant supplemental information (Section 4.2.2) (known as title-abstract screening).

Studies that were considered relevant to the toxicity assessment after title-abstract screening advanced to a second round of automated prioritization. Scientists then manually reviewed the full texts of the prioritized studies (Section 4.2.5) against a stricter set of inclusion criteria to identify human studies of neurodevelopmental or dental fluorosis health effects (i.e., studies relevant to the specific scope of the toxicity assessment).

4.2.1. Initial PECO Criteria

The initial PECO criteria (Table 4-1) were developed to guide title-abstract screening and define the conditions that studies must meet to be considered further for potential inclusion in the toxicity assessment. To meet the criteria, a study must meet all PECO elements of the criteria. The initial PECO criteria were developed during problem formulation and are intentionally broad for the purpose of identifying all potentially relevant evidence, including both human and animal studies.

Table 4-1. Initial PECO Criteria for the Fluoride Human Health Toxicity Assessment[a]

PECO Element	Inclusion Criteria[b]
Populations	Human: Any population and lifestage (occupational or general population, including children and other sensitive populations). Animal: Nonhuman mammalian animal species (whole organism) of any lifestage (including preconception, *in utero*, lactation, peripubertal, and adult stages).
Exposures	Relevant forms: Any exposure to fluoride and/or related forms, including but not limited to: Fluoride (CASRN 16984-48-8). Other names/forms: ammonium fluoride calcium fluoride disodium hexafluorosilicate fluoride ion fluorine fluorosilic acid hydrofluoric acid, ion(1-) hydrofluorosilicate hydrofluosilicic acid perfluoride potassium fluoride sodium hexafluorosilicate sodium fluoride sodium fluorosilicate sodium silicofluoride amine fluoride olaflur The following forms should also be included if mentioned in the context of exposure via toothpaste: sodium monofluorophosphate, stannous fluoride, amine fluoride, and acidulated phosphate fluoride. Aluminum fluoride will be excluded due to complex interactions with fluoride. All PFAS and fluoridated organic compounds (e.g., fluoridated pyrimidines) will be excluded. Human: Any exposure to the chemicals listed via oral routes or unknown/multiple routes (e.g., biomonitoring studies). Other exposure routes, including inhalation or dermal, will be tagged as supplemental. Animal: Any exposure to the chemicals listed via oral routes. Other exposure routes, including inhalation, dermal, injection, or unknown/multiple routes will be tagged as supplemental. Studies involving exposures to mixtures will be included only if they include exposure to the chemical of interest alone. Studies with fewer than 28 days of dosing, with the exception of reproductive, developmental, immune, and neurological health outcome studies, will be tagged as supplemental. Animal studies with only one dose group (e.g., control and one dose group only) will be tagged as supplemental.
Comparators	Human: A comparison or referent population exposed to lower levels (or no exposure/exposure below detection limits) of the chemical or interest, or exposure to the chemical of interest for shorter periods of time. Case reports and case series will not be considered PECO-relevant but will be tagged as supplemental if other PECO criteria are met. Animal: A concurrent control group exposed to vehicle-only treatment or untreated control. Controls with low levels of fluoride in drinking water (tap water) and animal feed should be included if other PECO parameters are met.
Outcomes	All health outcomes (both cancer and noncancer).
PBPK Models	Studies describing PK or PBPK models will be included. Classical PK or Dosimetry Model Studies: Classical PK or dosimetry modeling usually divides the body into just one or two compartments, which are not specified by physiology, where movement of a chemical into, between, and out of the compartments is quantified empirically by fitting model

PECO Element	Inclusion Criteria[b]
	parameters to ADME data. This category is for papers that provide detailed descriptions of PK models that are not a PBPK model. Note: ADME studies often report classical PK parameters, such as bioavailability (fraction of an oral dose absorbed), volume of distribution, clearance rate, or half-life. If a paper only provides such results in tables with minimal description of the underlying model or software (i.e., uses standard PK software without elaboration), including "noncompartmental analysis," it should be listed only as a supplemental material ADME study. PBPK or Mechanistic Dosimetry Model Studies: PBPK models represent the body as various compartments (e.g., liver, lung, slowly perfused tissue, richly perfused tissue) to quantify the movement of chemicals or particles into and out of the body (compartments) by defined routes of exposure, metabolism and elimination, and thereby estimate concentrations in blood or target tissues.

PECO = populations, exposures, comparators, and outcomes; CASRN = Chemical Abstracts Service registry number; PFAS = per- and polyfluoroalkyl substances; PBPK = physiologically based pharmacokinetic; PK = pharmacokinetic; ADME = absorption, distribution, metabolism, and excretion.

[a]The initial PECO criteria are used during title-abstract screening.

[b]In addition to studies meeting the initial PECO criteria, mechanistic studies (including in vitro studies) are tracked as 'included' during title-abstract screening to ensure that they will be included by the machine-learning model in SWIFT-Active Screener software during the prioritization step that takes place prior to full-text review (Section 4.2.3.1). Mechanistic studies are tracked as 'potentially relevant supplemental material' during full-text review.

4.2.2. Potentially Relevant Supplemental Material

Studies containing supplemental material potentially relevant to the toxicity assessment are inventoried during the literature screening process. Such studies are categorized ("tagged") into the different categories described in Table 4-2 for potential subsequent review at later stages of the toxicity assessment development as appropriate or as needed. As toxicity assessment development progresses, studies containing potentially relevant supplemental material could emerge as being critically important to the toxicity assessment and need to be evaluated (e.g., ADME studies), might be helpful for providing context, or might not be cited at all in the toxicity assessment.

Table 4-2. Categories of Potentially Relevant Supplemental Material for the Fluoride Human Health Toxicity Assessment

Category	Evidence
Mechanistic studies	Studies reporting measurements related to a health outcome that inform the biological or chemical events associated with phenotypic effects, in both mammalian and non-mammalian (genotoxicity) model systems, including in vitro, in vivo (by various routes of exposure), ex vivo, and in silico studies.
Beneficial-only health outcome	Studies reporting only the beneficial effects of fluoride treatment or exposure without quantitatively reporting adverse effects or side effects. Beneficial is defined based on the intent of the study (e.g., was the study designed to assess the use of fluoride to prevent cavities). Studies reporting an adverse outcome with a potentially beneficial response (e.g., observational study reporting reduced risk of cancer) will not be considered beneficial-only studies and will be included as PECO-relevant.
Non-mammalian model systems	Studies in non-mammalian model systems (e.g., fish, birds, *C. elegans*).

Category	Evidence
ADME	Studies providing information regarding ADME.
Acute/short-term duration exposures	Animal studies of fewer than 28 days (unless the study is a developmental and reproductive health outcome study).
Only one exposure group	Animal studies with only one exposure group (e.g., control and 1 mg/kg/day fluoride).
Non-oral routes of exposure	Studies addressing routes of exposure that fall outside the PECO scope, including inhalation and dermal exposure routes.
Exposure characteristics (no health outcome)	Exposure characteristics studies include data that are unrelated to toxicological endpoints, but which provide information on exposure sources or measurement properties of the environmental agent (e.g., demonstrate a biomarker of exposure).
Susceptible populations (no health outcome)	Studies that identify potentially susceptible subgroups (e.g., studies that focus on a specific demographic, lifestage, or genotype).
Environmental fate or occurrence (including food)	Studies that focus on describing where the chemical will end up after it is used and released into the environment.
Mixture studies	Mixture studies that are not considered PECO-relevant because they do not contain an exposure or treatment group assessing only the chemical of interest.
Case studies or case series	Case reports and case series will be tracked as potentially relevant supplemental information if other PECO parameters are met.
Records with no original data	Records that do not contain original/primary data, such as other agency assessments, informative scientific literature reviews, meta-analyses, editorials, or commentaries.
Posters or conference abstracts	Records that do not contain sufficient documentation to support study evaluation and data extraction.

PECO = populations, exposures, comparators, and outcomes; ADME = absorption, distribution, metabolism, and excretion.

4.2.3. Automated Prioritization Approaches

Prior to manual review of studies by subject matter experts, a variety of natural language processing and machine-learning techniques can be applied to large evidence bases to efficiently identify studies most likely to be relevant. The automated tools are well-established for the systematic review of health effects literature and have been used to develop a number of final EPA toxicity assessments (e.g., U.S. EPA (2025e, 2024e, 2023c, 2019c)). Some of these techniques are applied as part of the literature screening approach for the future Fluoride Human Health Toxicity Assessment and are briefly described in the following subsections.

A more detailed description of how these methods were applied, including the EPA's manual checks of the automated prioritization results, is provided in Appendix C.2.

4.2.3.1. SWIFT-Review Filtering

Sciome's SWIFT-Review literature prioritization software has preset literature search strategies ("filters" or "tags") developed by information specialists that can be used to identify studies most likely to contain relevant human health effects information. The filters function like a typical search strategy in which studies are tagged or categorized as belonging to a certain filter if the terms in the filter literature search strategy appear in title, abstract, keyword, or MeSH fields of the study. The details of the search strategies that underlie the filters are available online. SWIFT-Review filtering is an established systematic review tool that is validated (Howard et al., 2016) and consistently used in EPA assessments (U.S. EPA, 2024a, b, 2022b).

4.2.3.1.1. Prioritization Prior to Title-Abstract Screening

Prior to title-abstract screening, studies are categorized under one or more of the following SWIFT-Review evidence stream (i.e., data type) filters to prioritized references that are most likely to meet the broad initial PECO criteria:

- Human
- Epidemiologic quantitative analyses
- Animal (human health models)
- Animal (all)
- In vitro

In addition, studies that SWIFT-Review is unable to categorize under an evidence stream ("no tag") are further investigated for potential relevance with topic extraction (Section 4.2.3.3). Studies that do not receive at least one of the evidence stream tags listed above (or no tag) are not considered further.

4.2.3.1.2. Prioritization Prior to Full-Text Review

Prior to full-text review, further SWIFT-Review filtering is applied to studies that advance from title-abstract screening to identify the studies most likely to be relevant to the toxicity assessment's scope (see Section 2.3). Studies that are tagged with at least one of the following health outcome tags are prioritized for full-text review:

- Nervous system
- Dental
- Cancer (included for systematic evidence map described in the assessment plan)

Studies are also tagged to evidence stream filters at this stage (using the same approach described for title-abstract screening prioritization) to prioritize human and animal studies for review by screeners with appropriate expertise.

The standard search strings contained in SWIFT-Review are applied for the evidence stream and cancer tags; however, the nervous system and dental tags were developed by the EPA specifically for the future Fluoride Human Health Toxicity Assessment (during the assessment plan), to help identify studies on neurodevelopmental and dental effects prior to full-text review against the refined PECO criteria (Section 4.2.5). The search string underlying the existing SWIFT-Review nervous system tag was refined by the EPA to specifically target neurodevelopmental effects, and a new dental effects search string and tag was developed by the

EPA (Appendix C.2.2.1). To validate the new strings, references from existing assessments with known health outcomes (e.g., studies on neurodevelopmental outcomes cited in the NTP monograph (NTP, 2024) and studies on dental effects cited in the 2008 EPA dose-response analysis (U.S. EPA, 2008b) were tested with SWIFT-Review. The titles of tagged studies were reviewed for accuracy, and updates were made to new strings to ensure all relevant health outcome references from existing assessments were appropriately tagged.

4.2.3.2. Supervised Clustering

Supervised clustering combines aspects of both unsupervised and supervised machine learning to classify documents based on a small training set of relevant literature (known as "seed studies") (Varghese et al., 2018). Supervised clustering has been used in other final EPA toxicity assessments (U.S. EPA, 2025e, 2019c) to sort large literature databases and efficiently identify studies most likely to meet PECO criteria. Supervised clustering was utilized prior to both title-abstract and full-text screening of the initial literature search results and is also used for updated literature search results. The details of how this approach is applied, including the EPA's manual checks of the clustering performed with the initial literature search results, are discussed in Appendix C.2.

For the future Fluoride Human Health Toxicity Assessment, ensemble supervised clustering in ICF's Litstream™ software is used. This software uses six clustering algorithms. If any one of the six models finds a study to be relevant, it is tagged as such (thus erring on the side of inclusion/retention). An ensemble score is generated by adding the results of the six models. A study with an ensemble score of six indicates that the study was found to be relevant by all six models. Studies are prioritized for screening by descending order of ensemble score (6 to 0), where a score of 6 indicates studies that are most likely to be relevant and a score of 0 indicates studies that are least likely to be relevant.

4.2.3.3. Topic Extraction

Topic extraction is performed using ICF's Document Classification and Topic Extraction Resource (DoCTER) (integrated in Litstream), which clusters studies into groups based on similarity of keywords used in the titles and abstracts. DoCTER assigns each reference to a single cluster and provides the most common topic keywords for each cluster. The topics are then reviewed by subject matter experts to identify whether they contain terminology that is expected to be relevant to the assessment. If none of the topics include relevant terminology, they are excluded. If potential relevance for a topic is unclear, titles and abstracts within the topic are manually reviewed for potential relevance. This process can be performed in an iterative manner to further narrow topics for which there are large numbers of studies. This approach was applied prior to title-abstract screening for the initial database searches and a detailed description of the results is provided in Appendix C.2.

4.2.4. Title-Abstract Screening Using Initial PECO Criteria

Studies that are prioritized by the initial round of automated tools undergo manual title-abstract screening where the title and abstract of each study is reviewed against the initial PECO criteria (Table 4-1). Screening is performed by two independent reviewers per study using Sciome's

SWIFT-Active Screener software (Howard et al., 2020). SWIFT-Active Screener is an established systematic review tool that is routinely used in the EPA's systematic review process (U.S. EPA, 2024a, b, 2022b). It has a machine-learning component that is used to predict relevant references and expedite the manual screening of large evidence bases by moving the references predicted to be relevant to the front of the queue so that they are screened first by reviewers. Studies are screened until the software's algorithm predicts that 95% of the relevant studies have been screened and included as relevant. Published simulations have shown that the use of a 95% threshold has actually resulted in inclusion of 99% of relevant studies (Howard et al., 2020). If results of the two independent reviews conflict, the conflict is resolved by a senior-level tertiary reviewer. Studies determined to meet the initial PECO criteria during title-abstract screening are designated as 'included' and advance to further automated prioritization (Section 4.2.3.1.2) and full-text review (Section 4.2.5). Studies that report mechanistic data (e.g., in vitro studies) are also included and tracked as 'potentially relevant supplemental material' during title-abstract screening to ensure that they will be prioritized by SWIFT-Active Screener's machine-learning model in the further automated prioritization step.

Following completion of title-abstract screening, additional review by a senior-level screener is conducted for 15% of excluded studies to ensure quality of the final title-abstract results and that relevant studies were not missed. If no relevant studies are identified in this selection of excluded studies, no further action is taken. If relevant studies are identified in the group of excluded studies, additional review is conducted for a larger portion of the excluded studies.

4.2.5. Full-Text Review Using Refined PECO Criteria, and Literature Inventory Development

The EPA developed a set of refined PECO criteria (Table 4-3) that are applied during full-text screening and reflect the final scope of toxicity assessment (Section 2.3). According to the refined PECO criteria, human studies that evaluate fluoride exposure and dental fluorosis or neurodevelopmental effects are considered relevant. PK or PBPK model studies are also included as relevant.

Studies that reach the full-text review step are manually screened using ICF's Litstream™ software. Litstream™ is a systematic review toolbox that includes collaborative literature screening as described below. Full-text copies of the studies are retrieved and stored in the EPA's Health and Environmental Research Online (HERO) platform. Detailed reference information is made publicly available on the HERO project page when the draft toxicity assessment is released.

Studies meeting the refined PECO criteria are tagged as 'included,' and key data from the studies (e.g., information on study design, study populations, exposure measurement information, health outcomes that were assessed, and whether the studies report quantitative data with sufficient detail to potentially support dose-response analysis) are extracted using Litstream™ to create a literature inventory.

As with title-abstract screening, studies that include potentially relevant supplemental information are tagged as such. Studies that meet the initial PECO criteria but do not meet the refined PECO criteria are screened and tagged but do not undergo literature inventory data

extraction, except for studies reporting on the development of cancer following fluoride exposure. A primary reviewer completes the initial full-text review (and literature inventory data extraction, if applicable), and all studies undergo secondary review for QA by a senior-level reviewer.

Table 4-3. Refined PECO Criteria for the Fluoride Human Health Toxicity Assessment

PECO Element	Inclusion Criteria
Populations	Human: Any population and lifestage.
Exposures	Relevant forms: Any exposure to fluoride, and/or related forms (list in initial PECO criteria, Table 4-1). Human: Any exposure to the chemicals listed above via oral routes or unknown/multiple routes (e.g., biomonitoring studies) occurring during lifestages ranging from the fetus through adolescence. Other exposure routes, including inhalation or dermal, will be tagged as supplemental.
Comparators	Human: A comparison or referent population exposed to lower levels (or no exposure/exposure below detection limits) of the chemical of interest, or exposure to the chemical of interest for shorter periods of time. Case reports and case series will be tagged as supplemental if other PECO criteria are met.
Outcomes	Human: Dental fluorosis and neurodevelopmental outcomes (including but not limited to cognition, motor, and behavior). Studies of dental caries alone without additional priority outcomes will not be included as they are primarily assessing beneficial effects of exposure.
PBPK Models	Studies describing PK or PBPK models will be included (initial PECO criteria for details, Table 4-1).

PECO = populations, exposures, comparators, and outcomes; PBPK = physiologically based pharmacokinetic; PK = pharmacokinetic.

4.2.6. Non-English Language Studies

Although the literature search strategy is designed to primarily capture English language studies, non-English language studies are evaluated for potential relevance. Non-English language studies are identified from the literature search results based on following criteria:

- Study receives a "Non-English reports" tag during manual title-abstract screening (i.e., screening against initial PECO criteria) and/or full-text review (i.e., screening against refined PECO criteria)
- Study is not manually screened but the reference for the study appears in SWIFT-Active Screener with brackets in the title and/or abstract field

These studies are further prioritized using SWIFT-Review filtering prior to title-abstract screening, since most non-English language studies have translated titles and abstracts. Studies tagged with the following evidence stream and health outcome tags are prioritized for further review:

- Human – Developmental neurotoxicity
- Human – Dental
- Human – Cancer
- Animal – Cancer

These tags were prioritized based off of the narrowed scope of the assessment and refined PECO (i.e., human studies that evaluate neurodevelopmental and dental effects). Human and animal

cancer tags were included so the EPA could develop the systematic evidence map described in the assessment plan (U.S. EPA, 2026). In addition to receiving a prioritized tag, a study must include both a title and an abstract to advance to title-abstract screening. Title-abstract screening is performed according to the approach described in Section 4.2.4. Full-text review largely follows the same approach, as outlined in Section 4.2.5. Native speakers review studies in the published language when possible; if native speakers are unavailable, then translated versions of each study are obtained by uploading the original study file to Google Translate.

4.3. Health Effects Literature Survey Results

Section Overview: This section summarizes the results of the initial literature search and the review of the search results to identify relevant health effects studies for the future Fluoride Human Health Toxicity Assessment. The EPA developed an interactive diagram that allows users to see how individual studies were identified and categorized during each step of the literature review.

Figure 4-1 shows a literature flow diagram for the health effects studies, which provides a visual overview of the process and results for the initial literature search and screening that was performed during the assessment plan (U.S. EPA, 2026). An online interactive version allows users to filter specific references. Additional information, including detailed descriptions of the automated prioritization approaches that were applied prior to manual title-abstract and full-text review, is provided in Appendix C.

The initial database searches were conducted on November 1, 2024, and identified 268,093 unique references. Automated prioritization approaches, including SWIFT-Review tagging, supervised clustering, and topic extraction, were then used to identify studies most likely to be informative for the toxicity assessment prior to manual title-abstract screening. A total of 194,696 references were excluded based on these approaches. Additionally, the crosswalk of reference lists from existing fluoride assessments and other sources (Section 4.1.2) with the literature search results identified an additional 297 potentially relevant studies, including 46 studies that had been captured in the database searches but excluded during automated prioritization processes. These 297 studies advanced directly to title-abstract screening. In total, 73,703 studies advanced to title-abstract screening.

During title-abstract screening, subject matter experts reviewed a total of 34,703 studies to reach the predicted 95% inclusion threshold for SWIFT-Active Screener. Of these, 6,675 studies were identified as PECO-relevant, and 14,282 were identified as containing potentially relevant supplemental material. A total of 52,746 were either excluded as not PECO-relevant during manual review, or were not prioritized for manual review (i.e., were excluded) by the machine-learning model. Roughly 2,400 of the studies that were excluded because they were not prioritized by SWIFT-Active Screener were manually reviewed by subject matter experts to evaluate the accuracy of SWIFT-Active Screener's machine-learning model. No inaccuracies were identified.

The 6,675 studies identified as PECO-relevant during title-abstract screening underwent prioritization with SWIFT-Review tagging and supervised clustering prior to full-text review (Section 4.2.3 and Appendix C.2.2). As a result of these prioritization approaches, 4,020 studies were excluded and 2,655 advanced to full-text review against a refined set of PECO criteria (Section 4.2.5). As a result of full-text review, 671 studies met PECO criteria, 978 were tagged as containing potentially relevant supplemental material, 979 were excluded, and 27 were included in the cancer literature inventory.

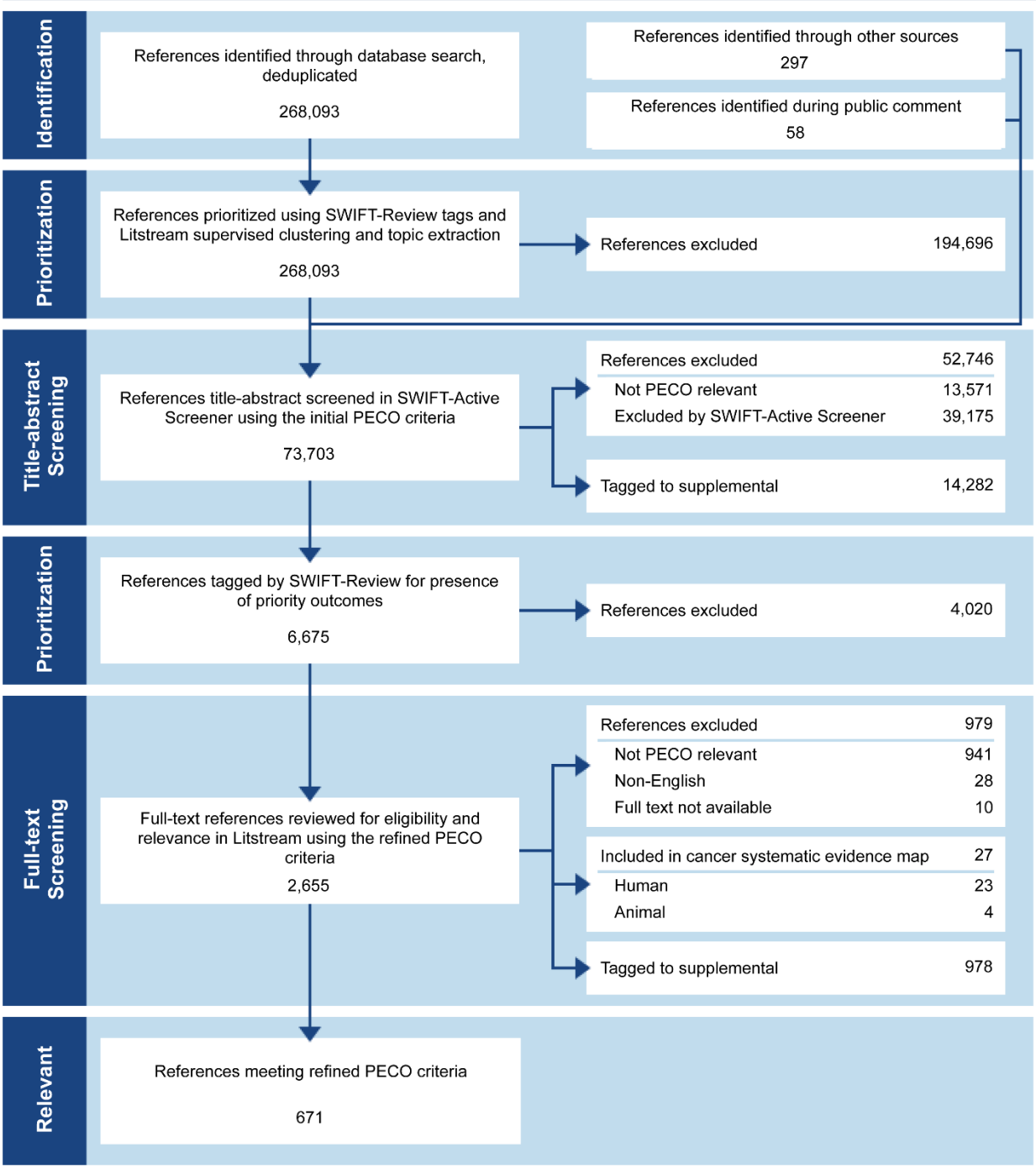


Figure 4-1. Literature Flow Diagram for the Health Effects Preliminary Literature Survey

The assessment plan included initial screening results through title-abstract screening. The screening results presented here have been updated to reflect the final results of the preliminary literature survey through full-text screening. The EPA plans to update the literature search prior to releasing the draft assessment.

[*] Eight additional references were identified from public comments on the assessment plan but were published after the literature search date cutoff. These references will be screened as part of the upcoming literature search update.

4.4. PK/PBPK Model Preliminary Literature Survey Results

Section Overview: This section summarizes the results of the initial literature search and screening with respect to PK and PBPK model studies. These studies are important for the toxicity assessment because they help EPA scientists estimate how fluoride is absorbed by, moves through, and is removed by the human body. This information will also help the EPA address key science issues, including how the agency can use measures of fluoride in the body (e.g., urine concentrations) to estimate actual fluoride intake. The EPA's process for reviewing and evaluating PK models for use in the toxicity assessment follows Gold Standard Science objectives by ensuring the agency identifies the best-available science that describes connection between biological measurements and the actual amount of fluoride consumed in a manner that is transparent and reproducible.

As noted in the initial PECO criteria outlined in Table 4-1, studies containing fluoride PK and PBPK models were included as part of the preliminary literature survey. This section provides a brief overview of the studies of PK/PBPK models identified through the initial literature search. Additional models describing specific components of fluoride ADME (e.g., bone uptake and oral clearance) are described in Appendix C.4. A description of how these models will be evaluated and considered in the toxicity assessment is provided in Section 6.3.

Two fluoride PBPK models designed to predict systemic and urinary fluoride concentrations in humans were identified. Table 4-4 provides a brief summary of the models. The original model (Rao et al., 1995) includes lung; liver; kidney; rapidly and slowly perfused bone; and surface and bulk bone compartments parameterized for rats and adult humans. Subsequently, Jean et al. (2018), using age-specific physiology equations from Price et al. (2003), developed a child-specific adaptation of Rao et al. (1995) to account for differences in urinary fluoride clearance and bone uptake between children and adults. However, this version only had a single bone compartment and did not include resorption from the bone to the blood.

Table 4-4. Studies Reporting Physiologically Based Pharmacokinetic Models for Fluoride

Source	Population and Exposure	Compartments	Internal Dose Metrics	Notes
(Rao et al., 1995)	Adults; life-long exposure	Liver Kidney Rapidly/slowly perfused, bone Bone subdivided into surface bone and inner bulk bone	Bone and blood fluoride concentration, renal/urinary elimination rate	Designed for long-term bone kinetics over lifetime exposure; detailed treatment of bone surface vs. bulk bone to simulate accumulation.
(Jean et al., 2018)	Children (4- and 8-year-olds); steady-state exposure	Liver Kidney Rapidly/slowly perfused, bone Bone simplified into single (combined surface/bulk) compartment	Blood fluoride and renal/urinary fluoride excretion at steady state	Used age-specific physiology equations from (Price et al., 2003); provides a "snapshot" of pharmacokinetics for a given age; predicts flux to bone but not amount or concentration in bone.

While PBPK models are preferred because they consider various physiological aspects within the body (U.S. EPA, 2014b), compartmental PK models, which represent the body using one-, two-, or three-compartment mathematical frameworks, remain informative for understanding how fluoride is absorbed and excreted. Following the review of the literature, six compartment-based fluoride PK models derived from observed human data were identified. Referred to as "classical pharmacokinetic models," these single- and multicompartment models provide PK parameters that can be used to assess internal body burden following exposure. These parameters include blood concentration area under the concentration curve (AUC), oral bioavailability, urinary excretion rate and renal clearance, systemic half-life, and extrarenal (e.g., bone) clearance. Although these parameters do not characterize the fine-grained fluoride pharmacokinetics within specific tissues, they do provide information amenable to converting measured fluoride concentrations in blood/urine to applied dose. Table 4-5 provides a brief summary of the identified fluoride PK models.

Table 4-5. Studies Reporting Compartmental Pharmacokinetic Models

Source	Population and Exposure	Compartments	Internal Dose Metrics	Notes
(Ekstrand and Ehrnebo, 1979)	Adults; oral sodium fluoride administration while fasting, with milk, or with breakfast	Single compartment	Time-course blood fluoride concentration, AUC, oral bioavailability	Intravenous administration used to calculate bioavailability
(Ekstrand et al., 1977)	Adults; oral sodium fluoride administration in both single- and multiple-dose regimens	Two- and three-compartment evaluation	Time-course blood fluoride concentrations from range of single and repeat sodium fluoride doses; AUC, blood C_{ss}, half-life; urinary excretion	Controlled dose experiments using both single- and repeat doses
(Spak et al., 1982)	Adults; oral fluoride in water, milk, or baby formula	Single compartment	Time-course blood fluoride concentration, AUC, urinary excretion rate, bioavailability, renal fluoride clearance	Decrease in bioavailability associated with fluoridated milk and baby formula
(Ekstrand and Ehrnebo, 1983)	Adults; 10 mg sodium fluoride	Single compartment	Time-course blood fluoride concentration, urinary excretion rate, blood/urine half-life	Plasma blood concentration highly correlated with urinary excretion rate; renal clearance depends on urine pH
(Kraenzlin et al., 1990)	Osteoporotic adults; 10 mg fluoride	Single compartment	Time-course blood fluoride concentration, plasma, renal, bone clearance	Extrarenal (bone) removal from blood characterized through one-compartment model by subtracting plasma fluoride clearance from renal clearance
(Ekstrand and Spak, 1990)	Adults; multiple sodium fluoride tablet formulations; co-exposure to calcium	Two compartments	Time-course blood fluoride concentration, C_{max}, T_{max}, AUC	Calcium-specific changes to fluoride pharmacokinetics

AUC = area under the concentration curve; C_{ss} = concentration at steady state; C_{max} = maximum blood concentration; T_{max} = time to reach maximum blood concentration.

5. Refined Evaluation Plan

Section Overview: This section describes additional factors used to identify the studies most likely to contain information that will be useful for determining the level of fluoride that a person can ingest daily and be unlikely to experience harmful health effects (i.e., toxicity value). These factors are used by EPA scientists to find the best available science in a manner that is transparent and reproducible. All studies are prioritized based on whether they: could be used to derive a toxicity value; measured fluoride exposure during an appropriate period; and considered potential confounding (i.e., factors that may distort the true results). Additional considerations that are specific to the selection of studies on neurodevelopmental effects and dental fluorosis (e.g., how the studies measured the exposure or outcome) are also described. Prioritized studies move on to study evaluation (Section 6), which is the next step of the systematic review process.

This section outlines how reviewers prioritize the set of PECO-relevant health effects studies most likely to be relevant to dose-response analysis. The identified studies then undergo study evaluation (Section 6), focusing the evaluation efforts on the studies that are most likely to be relevant for dose-response analysis (Section 8). Information in the studies identified through literature screening and the development of the literature inventory (e.g., the types of exposure measures used in the studies) is used to determine which studies move forward to study evaluation. The factors described below are applied to initially identify potential dose-response studies, but each factor is further elaborated on in Section 6 and comprehensively considered during study evaluation. Reviewers err on the side of inclusion if there is uncertainty about whether information is sufficient.

Reviewers assess whether a study could be sufficient for dose-response analysis (i.e., tagged "to be assessed further" in the inventory) based on the following factors:

- The study reports quantitative data or effect estimates that appear suitable for dose-response analysis.
- The study uses a measure of fluoride exposure that represents exposure during an etiologically relevant period (the time frame when an exposure could plausibly cause the outcome). For example, exposure data could be collected prospectively during an etiologically relevant period or might be collected concurrent with the outcome but can reasonably be considered representative of exposure during an etiologically relevant period.
- The study includes some consideration of potential confounding. Confounding can be considered through study design (e.g., restriction of study population) or during statistical analysis (e.g., restriction, stratification, or formal adjustment). Therefore, lack of formal control for confounding may be acceptable in some studies (e.g., studies of dental fluorosis with populations restricted based on age).

Studies of neurodevelopmental effects should additionally either:

- Measure exposure on an individual basis rather than at the population level, since this allows for more precise exposure quantification and reduces concern with exposure misclassification; or
- Use a population-level measure of fluoride (e.g., drinking water measures at the municipality level) but incorporate individual information to estimate exposure (e.g., drinking water ingestion rates to estimate individual intake).

Studies of dental fluorosis should additionally:

- Be conducted in a population experiencing fluoride exposure at concentrations that may be relevant to populations in the United States (e.g., majority of the study sample experiencing fluoride concentrations in drinking water that are at or below 4 mg/L).
- Determine the presence and severity of dental fluorosis using a validated instrument for measuring fluorosis, such as the Dean's Index, Thylstrup-Fejerskov Index (TFI), or the Tooth Surface Index of Fluorosis (TSIF).
- Report dental fluorosis prevalence or incidence at a person-level and not at the level of the tooth or tooth surface. Person-level results (e.g., the number and proportion of people with dental fluorosis) are more useful in evaluating the occurrence of adverse fluorosis in the population compared with results for individual teeth or tooth surfaces.
- Measure dental fluorosis on permanent (e.g., adult) rather than deciduous (e.g., baby) teeth. Effects on deciduous teeth are considered temporary.
- Report cases of moderate and/or severe dental fluorosis and include an analysis of moderate and/or severe dental fluorosis (e.g., as opposed to an analysis of any form of dental fluorosis). While there is continued debate about the adversity of different levels of dental fluorosis, the future Fluoride Human Health Toxicity Assessment does not plan to consider very mild or mild dental fluorosis as outcomes for RfD derivation.

6. Study Evaluation Strategy

Section Overview: This section describes the study evaluation approach for the future Fluoride Human Health Toxicity Assessment. Study evaluation is the process that the EPA uses to identify the best-available science by assessing the strengths and limitations of studies.

The general approach for human health effects studies is described in Section 6.1. Studies are assessed for potential sources of bias (i.e., anything that could affect the reliability of study results) and sensitivity (i.e., the ability of a study to detect true effects). Studies are evaluated across different categories (domains) based on specific criteria, and then overall confidence in each study is determined. Different findings or analyses within a single study can have different evaluation ratings.

Section 6.2 describes the specific criteria that are used to evaluate the human health effects studies on fluoride. Some of the criteria used in this process are customized to both fluoride exposure (e.g., what methods were used to evaluate fluoride exposure, and are they reliable?) and the outcomes that are the focus of this assessment (e.g., what methods were used to evaluate neurodevelopmental effects or dental fluorosis, and are they reliable?). These criteria help ensure that the EPA consistently and transparently evaluates and documents study features that are important for addressing some of the key science issues for the toxicity assessment (Section 2.4), such as the validity of exposure measurements in each study.

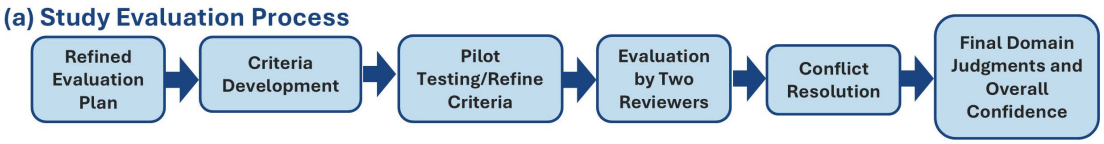
Section 6.3 describes the EPA's approach for using fluoride pharmacokinetic (PK) models in the future Fluoride Human Health Toxicity Assessment. PK models are sets of mathematical equations that estimate how fluoride is absorbed by, moves through, and is removed by the human body and help determine the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (i.e., the reference dose or RfD). PK models use measurements of fluoride exposure (e.g., fluoride concentration in blood or urine) to estimate an individual's total fluoride exposure over a lifetime. This section also outlines the evaluation approach for existing PK models that were identified in the literature search (Section 4.4).

6.1. Overview of Study Evaluation Approach for Human Health Effects Studies Applicable to the Fluoride Human Health Toxicity Assessment

Two key concerns for the evaluation of epidemiology studies in general are risk of bias, which is the assessment of internal validity (i.e., factors that affect the magnitude or direction of an effect in either direction), and sensitivity (i.e., factors that affect the ability of a study to detect a true effect and/or the ability to evaluate dose-response relationships). Reporting quality (e.g., completeness of reporting in a study, and whether there is selective reporting) is evaluated to determine the extent to which the information provided in each study allows for assessing these considerations. Study evaluations are aimed at discerning the expected magnitude and direction of any identified biases or other limitations (i.e., focusing on those that could substantively change a result). Conflict of interest is not explicitly assessed because evaluations of risk of bias

and study sensitivity are designed to encompass the primary aspects of methodological design that could cause concern, irrespective of the sponsoring entity.

Figure 6-1 provides an overview of the EPA's study evaluation process for epidemiology studies, including the different aspects of a study that are typically evaluated (domains), the judgments that can be applied to each domain, and the approach for reaching an overall confidence judgment for a study (or for a specific analysis or outcome within a study). The framework for this process is consistent with that outlined in (U.S. EPA, 2022b).



(b) Evaluation Domains, Domain Judgments, Overall Confidence

Evaluation Domains (Epidemiology)
Participant Selection
Exposure Measurement
Outcome Ascertainment
Confounding
Analysis
Sensitivity
Selective Reporting

Domain Judgments	
++ Good	Appropriate study conduct relating to the domain and minor deficiencies not expected to influence results.
+ Adequate	A study that may have some limitations relating to the domain, but they are not likely to be severe or to have a notable impact on results.
- Deficient	Identified biases or deficiencies interpreted as likely to have had a notable impact on the results or prevent reliable interpretation of study findings.
-- Critically Deficient	A serious flaw identified that makes the observed effect(s) uninterpretable. Studies with a critical deficiency will almost always be considered overall *Uninformative*.

Overall Confidence	
High	No notable deficiencies or concerns identified; potential for bias unlikely or minimal; sensitive methodology.
Medium	Possible deficiencies or concerns noted, but resulting bias or lack of sensitivity would be unlikely to be of a notable degree.
Low	Deficiencies or concerns were noted, and the potential for substantive bias or inadequate sensitivity could have a significant impact on the study results or their interpretation.
Uninformative	Serious flaw(s) makes study results unusable for hazard identification or dose response.

Figure 6-1. Overview of General Study Evaluation Process for Epidemiological Studies: (a) Overview of the Evaluation Process; (b) Evaluation Domains, Domain Judgments, and Overall Confidence Judgments

Adapted from U.S. EPA (2022b).

Prior to initiating the study evaluation process, assessment- and/or chemical-specific criteria for evaluating studies are developed in consultation with technical experts (as needed) and consistent with relevant agency frameworks and guidance documents (e.g., agency guidance for neurotoxicity and developmental toxicity risk assessments (U.S. EPA, 1998, 1991). While developing these assessment- and chemical-specific criteria, a pilot phase is used to assess and

refine the study evaluation process, during which evaluations are compared and a consensus is reached between reviewers through consultations among the toxicity assessment team. Additional chemical-specific knowledge or methodological issues might emerge during the pilot phase and can prompt refinements to the study evaluation process. Refinements made during the pilot phase and subsequent implementation are documented as updates to the protocol.

For study evaluation, at least two independent team members review each study.[3] For studies that assess more than one exposure-outcome relationship, each relationship is evaluated separately because the utility of a study can vary for different exposure and outcome measures (e.g., for the Outcome Ascertainment domain, a study may achieve a *Good* rating for measurement of neurodevelopmental endpoints but a *Deficient* rating for measurement of dental fluorosis due to methodological issues, such as validity of methodology or reporting quality).

The EPA might ask study authors for missing or additional information that could affect the study evaluation outcome or would be required to conduct dose-response analysis (e.g., missing group size or variance descriptors such as standard deviation or confidence interval). The decision to seek missing information from study authors involves consideration of the additional information that would be useful, specifically with respect to any information that could result in a reevaluation of overall study confidence. Outreach to study authors is documented and considered unsuccessful if authors do not respond to an email or phone request within one month of the attempt to contact.

For each outcome in a study, reviewers reach a consensus judgment of *Good*, *Adequate*, *Deficient*, *Not Reported*, or *Critically Deficient* for each evaluation domain, and provide a rationale for each judgment (in some cases, studies may be given a judgment of *Not Applicable* for a domain; Section 6.2.3 provides an example). If a consensus is not reached through discussion, a third reviewer performs an independent review of the study and resolves the conflict. In addition to the study evaluation domains described in Figure 6-1, the evaluation will also consider each study's potential utility for dose-response analysis at this stage.

The following are the possible judgments that can be applied to each evaluation domain for each study:

- *Good* represents a judgment that the study was conducted appropriately in relation to the evaluation domain, and any deficiencies, if present, are minor and would not be expected to influence the study results.
- *Adequate* indicates a judgment that there are methodological limitations relating to the evaluation domain, but that those limitations are not likely to be severe or have a notable impact on the results.
- *Deficient* denotes that the study evaluation identified biases or deficiencies that are interpreted as likely to have a notable impact on the results or that might prevent reliable interpretation of the study findings.

[3]"Study" is used instead of a more accurate term (e.g., "experiment") throughout these sections, owing to an established familiarity within the field for discussing a study's risk of bias or sensitivity. However, all evaluations discussed herein are explicitly conducted at the level of an individual outcome within a group of humans with varying exposure levels.

- *Not Reported* indicates that the information necessary to evaluate the domain is not available in the study. Generally, this term carries the same functional interpretation as *Deficient* for the purposes of the study confidence classification (described later). Depending on the number and severity of other limitations identified in the study, it might or might not be worth reaching out to the study authors to obtain the missing information (as referenced earlier).
- *Critically Deficient* reflects a judgment that the study conduct introduces a serious flaw that makes the study uninterpretable. Studies with a determination of *Critically Deficient* for an evaluation domain are almost always considered *Uninformative* overall and thus not used for hazard identification or dose-response analysis, but they could be used to highlight potential research gaps. Given this potential for exclusion, this classification is used infrequently and with extreme care; methodological limitations warranting this classification are defined a priori on an exposure- and outcome-specific basis and are inherently severe enough to warrant exclusion based on a single critical deficiency.

Once a study is evaluated across all domains, the identified strengths and limitations are considered to determine overall study confidence of *High*, *Medium*, *Low*, or *Uninformative* for each exposure-outcome relationship within the study. Overall study confidence is based on the reviewer judgments across the evaluation domains and includes consideration of the likely impacts of any noted deficiencies with respect to bias and sensitivity. There are no predefined weights for the domains, and the reviewers apply expert judgment to determine the impacts of identified limitations on overall study confidence for a given outcome. The rationale for overall study confidence is documented clearly and consistently. This rationale should, to the extent possible, reflect an interpretation of the potential influence of study deficiencies on the results (including the direction and/or magnitude of influence).

The overall confidence classifications are as follows:

- *High* confidence: A well-conducted study with no notable deficiencies or concerns identified; the potential for bias is unlikely or minimal; and the study used sensitive methodology. *High*-confidence studies generally reflect judgments of *Good* across all or most evaluation domains.
- *Medium* confidence: A satisfactory (acceptable) study in which deficiencies or concerns are noted, but the limitations are unlikely to be consequential. Generally, *Medium*-confidence studies include *Adequate* or *Good* judgments across most domains, with the impact of any identified limitation not being judged as severe.
- *Low* confidence: A substandard study in which deficiencies or concerns are noted, and the potential for bias or inadequate sensitivity could have a substantial impact on the study results or their interpretation. Typically, *Low*-confidence studies have a *Deficient* evaluation for one or more domains, although some *Medium*-confidence studies might have a *Deficient* rating for domain(s) considered to have less influence on the magnitude or direction of effect estimates. *Low*-confidence studies are generally not used as the primary sources of information for derivation of toxicity values unless they are the only studies available. Studies rated as *Low* confidence only because of sensitivity concerns that result in a bias toward the null require additional consideration. Observing an effect

in these studies might increase confidence, assuming the study is otherwise well conducted.

- *Uninformative*: An unacceptable study in which serious flaw(s) make the study results unusable for informing dose-response analysis. Studies with *Critically Deficient* judgments in any evaluation domain are almost always classified as *Uninformative* (as explained earlier). Studies with multiple *Deficient* judgments might also be considered *Uninformative*. *Uninformative* studies are not considered further for dose-response analysis but might be useful for identifying possible research needs in the toxicity assessment conclusions.

Individual domain judgments and overall study confidence determined by each reviewer as well as the consensus judgments of reviewers are transparently documented in the EPA's version of the Health Assessment Workspace Collaborative (HAWC), a free and open-source, web-based software application.[4] Final study evaluations are housed in HAWC, including the rationales supporting the individual domain judgments and overall study confidence, are made available when the draft toxicity assessment is released to the public. Overall confidence determinations and their rationales are considered during selection of studies for dose-response analysis (Section 8.1).

6.2. Epidemiology Study Evaluation

During study evaluation, the epidemiology studies are evaluated for the following domains: exposure measurement, outcome ascertainment, participant selection, potential confounding, analysis, study sensitivity, and selective reporting. Table 6-1 presents the generic core, prompting, and follow-up questions intended to guide evaluation of each domain and apply to all exposures and outcomes. Core questions represent key concepts, whereas prompting and follow-up questions helped reviewers focus on relevant details within each domain. Table 6-1 also includes considerations that apply to all exposures and outcomes. Sections 6.2.1 to 6.2.3 describe fluoride-specific exposure measurement- and outcome-specific study evaluation criteria, in addition to the general criteria. These criteria are intended to improve consistency across reviewers by highlighting common issues specific to the future Fluoride Human Health Toxicity Assessment. They are used in conjunction with the generic considerations and not as a replacement.

[4]HAWC is a modular web-based interface to facilitate development of human health assessments.

Table 6-1. Questions and Considerations to Guide Study Evaluation for Each Domain in Epidemiology Studies[a]

Domain: Exposure Measurement

Core Question: Does the exposure measure reliably distinguish between different levels of exposure in a time window considered most relevant for a causal effect with respect to the outcome observation?

Prompting Questions	Follow-Up Questions
For all study types: Does the exposure measure capture the variability in exposure among the participants, considering intensity, frequency, and duration of exposure? Does the exposure measure reflect a relevant time window? If not, can the relationship between measures in this time and the relevant time window be estimated reliably? Is the exposure measurement likely to be affected by a knowledge of the outcome? Is the exposure measurement likely to be affected by the presence of the outcome (i.e., reverse causality)? ***For case-control studies of occupational exposures:*** Is exposure based on a comprehensive job history describing tasks, setting, time period, and use of specific materials? ***For biomarkers of exposure, general population:*** Is a standard assay used? What are the intra- and inter-assay coefficients of variation? Is the assay likely to be affected by contamination? Are values less than the limit of detection dealt with adequately? What exposure time period is reflected by the biomarker? If the half-life is short, what is the correlation between serial measurements of exposure?	Is the degree of exposure misclassification likely to vary by exposure level? If the correlation between exposure measurements is moderate, is there an adequate statistical approach to ameliorate variability in measurements? If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?

Considerations That Apply to Most Exposures and Outcomes

These considerations require customization to the outcome.

Good: Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. Exposure misclassification is expected to be minimal.

Adequate: Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. Exposure misclassification may exist but is not expected to greatly change the effect estimate.

Deficient:

Valid exposure assessment methods used, which represent the etiologically relevant time period of interest. Specific knowledge about the exposure and outcome raise concerns about reverse causality, but there is uncertainty whether it is influencing the effect estimate.

Exposed groups are expected to contain a notable proportion of unexposed or minimally exposed individuals, the method did not capture important temporal or spatial variation, or there is other evidence of exposure misclassification that would be expected to notably change the effect estimate.

Critically Deficient:

Exposure measurement does not characterize the etiologically relevant time period of exposure or is not valid.

There is evidence that reverse causality is very likely to account for the observed association.

Exposure measurement was not independent of outcome status.

Domain: Outcome Ascertainment	
Core Question: Does the outcome measure reliably distinguish between the presence or absence (or degree of severity) of the outcome?	
Prompting Questions	**Follow-Up Questions:**
For all study types: Is outcome ascertainment likely to be affected by knowledge of, or presence of, exposure (e.g., consider access to health care, if based on self-reported history of diagnosis)? *For case-control studies:* Is the comparison group without the outcome (e.g., controls in a case-control study) based on objective criteria with little or no likelihood of inclusion of people with the disease? ***For mortality measures:*** How well does cause of death data reflect occurrence of the disease in an individual? How well do mortality data reflect incidence of the disease? *For diagnosis of disease measures:* Is the diagnosis based on standard clinical criteria? If it is based on self-report of the diagnosis, what is the validity of this measure? ***For laboratory-based measures (e.g., hormone levels):*** Is a standard assay used? Does the assay have an acceptable level of inter-assay variability? Is the sensitivity of the assay appropriate for the outcome measure in this study population?	Is there a concern that any outcome misclassification is nondifferential, differential, or both? What is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?

Considerations That Apply to Most Exposures and Outcomes

These considerations require customization to the outcome.

Good:

High certainty in the outcome definition (i.e., specificity and sensitivity), minimal concerns with respect to misclassification.

Assessment instrument is validated in a population comparable to the one from which the study group was selected.

Adequate:

Moderate confidence that outcome definition was specific and sensitive, some uncertainty with respect to misclassification but not expected to greatly change the effect estimate.

Assessment instrument is validated but not necessarily in a population comparable to the study group.

Deficient:

Outcome definition was not specific or sensitive.

Uncertainty regarding validity of assessment instrument.

Critically Deficient:

Invalid/insensitive marker of outcome.

Outcome ascertainment is very likely to be affected by knowledge of, or presence of, exposure.

Note: Lack of blinding should not be automatically construed to be *Critically Deficient*.

Domain: Participant Selection	
Core Question: Is there evidence that selection into or out of the study (or analysis sample) is jointly related to exposure and to outcome?	
Prompting Questions	**Follow-Up Questions**
For longitudinal cohort: Did participants volunteer for the cohort based on knowledge of exposure and/or preclinical disease symptoms? Was entry into the cohort or continuation in the cohort related to exposure and outcome? ***For occupational cohort:*** Did entry into the cohort begin with the start of the exposure? Was follow-up or outcome assessment incomplete, and if so, was follow-up related to both exposure and outcome status? Could exposure produce symptoms that would result in a change in work assignment/work status ("healthy worker survivor effect")? ***For case-control study:*** Were controls representative of population and time periods from which cases were drawn? Are hospital controls selected from a group whose reason for admission is independent of exposure? Could recruitment strategies, eligibility criteria, or participation rates result in differential participation relating to both disease and exposure? ***For population-based survey:*** Was recruitment based on advertisement to people with knowledge of exposure, outcome, and hypothesis?	Are differences in participant enrollment and follow-up evaluated to assess bias? If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)? Are appropriate analyses performed to address changing exposures over time in relation to symptoms? Is there a comparison of participants and nonparticipants to address whether differential selection is likely?
Considerations that Apply to Most Exposures and Outcomes These considerations may require customization to the outcome. This could include determining what study designs effectively allow analyses of associations appropriate to the outcome measures (e.g., design to capture incident vs. prevalent cases, design to capture early pregnancy loss). **Good:** Minimal concern for selection bias based on description of recruitment process and follow-up (e.g., selection of comparison population, population-based random sample selection, recruitment from sampling frame including current and previous employees). Exclusion and inclusion criteria for participants specified and would not induce bias. Participation rate is reported at all steps of study (e.g., initial enrollment, follow-up, selection into analysis sample). If rate is not high, there is appropriate rationale for why it is unlikely to be related to exposure (e.g., comparison between participants and nonparticipants or other available information indicates differential selection is not likely). **Adequate:** Enough of a description of the recruitment process to be comfortable that there is no serious risk of bias. Inclusion and exclusion criteria for participants specified and would not induce bias. Participation rate is incompletely reported but available information indicates participation is unlikely to be related to exposure.	

Deficient: Little information on recruitment process, selection strategy, sampling framework or participation OR aspects of these processes raises the potential for bias (e.g., healthy worker effect, survivor bias). **Critically deficient:** Aspects of the processes for recruitment, selection strategy, sampling framework, or participation result in concern that selection bias is likely to have had a large impact on effect estimates (e.g., convenience sample with no information about recruitment and selection, cases and controls are recruited from different sources with different likelihood of exposure, recruitment materials stated outcome of interest and potential participants are aware of or are concerned about specific exposures).

Domain: Confounding	
Core Question: Is confounding of the effect of the exposure likely?	
Prompting Questions	**Follow-Up Question**
Is confounding adequately addressed by considerations in: Participant selection (matching or restriction)? Accurate information on potential confounders and statistical adjustment procedures? Lack of association between confounder and outcome, or confounder and exposure in the study? Information from other sources? Is the assessment of confounders based on a thoughtful review of published literature, potential relationships (e.g., as can be gained through directed acyclic graphing), and minimizing potential overcontrol (e.g., inclusion of a variable on the pathway between exposure and outcome)?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?

Considerations that Apply to Most Exposures and Outcomes

These considerations require customization to the exposure and outcome, but this may be limited to identifying key confounders.

Good:

Conveys strategy for identifying key confounders. This might include a priori biological considerations, published literature, causal diagrams, or statistical analyses; with recognition that not all "risk factors" are confounders.

Inclusion of potential confounders in statistical models not based solely on statistical significance criteria (e.g., $p < 0.05$ from stepwise regression).

Does not include variables in the models that are likely to be influential colliders or intermediates on the causal pathway.

Key confounders are evaluated appropriately and considered to be unlikely sources of substantial confounding. This often will include:

- Presenting the distribution of potential confounders by levels of the exposure of interest and/or the outcomes of interest (with amount of missing data noted);
- Consideration that potential confounders are rare among the study population or are expected to be poorly correlated with exposure of interest;
- Consideration of the most relevant functional forms of potential confounders;
- Examination of the potential impact of measurement error or missing data on confounder adjustment.
- Presenting a progression of model results with adjustments for different potential confounders, if warranted.

Adequate: Similar to **Good** but might not have included all key confounders, or less detail might be available on the evaluation of confounders (e.g., sub-bullets in **Good**). It is possible that residual confounding could explain part of the observed effect, but concern is minimal.

Deficient:

Does not include variables in the models that are likely to be influential colliders or intermediates on the causal pathway.

And any of the following:

The potential for bias to explain some of the results is high based on an inability to rule out residual confounding, such as a lack of demonstration that key confounders of the exposure-outcome relationships are considered;

Descriptive information on key confounders (e.g., their relationship relative to the outcomes and exposure levels) are not presented; or

Strategy of evaluating confounding is unclear or is not recommended (e.g., only based on statistical significance criteria or stepwise regression [forward or backward elimination]).

Critically Deficient: Includes variables in the models that are colliders and/or intermediates in the causal pathway, indicating that substantial bias is likely from this adjustment; or confounding is likely present and not accounted for, indicating that all of the results are most likely due to bias.

Domain: Analysis	
Core Question: Does the analysis strategy and presentation convey the necessary familiarity with the data and assumptions?	
Prompting Questions	**Follow-Up Question**
Are missing outcome, exposure, and confounder data recognized, and if necessary, accounted for in the analysis? Does the analysis appropriately consider variable distributions and modeling assumptions? Does the analysis appropriately consider subgroups of interest (e.g., based on variability in exposure level or duration or susceptibility)? Is an appropriate analysis used for the study design? Is effect modification considered, based on considerations developed a priori? Does the study include additional analyses addressing potential biases or limitations (i.e., sensitivity analyses)?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?

Considerations that Apply to Most Exposures and Outcomes

These considerations might require customization to the outcome. This could include the optimal characterization of the outcome variable and ideal statistical test (e.g., Cox regression).

Good:

Use of an optimal characterization of the outcome variable.

Quantitative results are presented (effect estimates and confidence limits or variability in estimates) (i.e., not presented only as a *p*-value or "significant"/"not significant").

Descriptive information about outcome and exposure is provided (where applicable).

Amount of missing data is noted and addressed appropriately (discussion of selection issues—missing at random vs. differential).

Where applicable, for exposure, includes the limit of detection (LOD), percentage below the LOD, and decision to use log transformation.

Includes analyses that address robustness of findings, e.g., examination of exposure-response (explicit consideration of nonlinear possibilities, quadratic, spline, or threshold/ceiling effects included, when feasible); relevant sensitivity analyses; effect modification examined based only on a priori rationale with sufficient numbers.

No deficiencies in analysis are evident. Discussion of some details might be absent (e.g., examination of outliers).

Adequate: Same as **Good**, except:

Descriptive information about exposure is provided (where applicable) but might be incomplete; might not have discussed missing data, cutpoints, or shape of distribution.

Includes analyses that address robustness of findings (examples in **Good**), but some important analyses are not performed.

Deficient:

Does not conduct analysis using optimal characterization of the outcome variable.

Descriptive information about exposure levels is not provided (where applicable).

Effect estimate and *p*-value are presented, without standard error or confidence interval.

Results are presented as statistically "significant"/"not significant."

Critically Deficient:

Results of analyses of effect modification are examined without clear a priori rationale and without providing main/principal effects (e.g., presentation only of statistically significant interactions that were not hypothesis driven).

Analysis methods are not appropriate for design or data of the study.

Domain: Selective Reporting	
Core Question: Is there reason to be concerned about selective reporting?	
Prompting Questions	**Follow-Up Question**
Are results provided for all the primary analyses described in the methods section? Is there appropriate justification for restricting the amount and type of results that are shown? Are only statistically significant results presented?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?

Considerations that Apply to Most Exposures and Outcomes

These considerations generally do not require customization and might have fewer than four levels.

Good: The results reported by study authors are consistent with the primary and secondary analyses described in a registered protocol or methods paper.

Adequate: The authors described their primary (and secondary) analyses in the methods section and results are reported for all primary analyses.

Deficient:

Concerns are raised based on previous publications, a methods paper, or a registered protocol indicating that analyses are planned or conducted that are not reported, or that hypotheses originally considered to be secondary are represented as primary in the reviewed paper.

Only subgroup analyses are reported, suggesting that results for the entire group are omitted.

Only statistically significant results are reported.

Domain: Sensitivity
Core Question: Is there a concern that sensitivity of the study is not adequate to detect an effect?
Prompting Questions
Is the exposure range adequate to detect associations and exposure-response relationships? Was the appropriate population included? Was the length of follow-up adequate? Is the time/age of outcome ascertainment optimal given the interval of exposure and the health outcome? Are there other aspects related to risk of bias or otherwise that raise concerns about sensitivity?
Considerations that Apply to Most Exposures and Outcomes These considerations might require customization to the exposure and outcome. Depending on the needs of the assessment, there might be fewer than four rating levels. Some study features that affect study sensitivity might have already been included in the other evaluation domains; these should be noted in this domain again, along with any features that have not been addressed elsewhere so that the rating provides an overall summary of factors that might impact sensitivity. When determining the overall study confidence rating, the evaluator should be conscious that a limitation could contribute to multiple domains and not double penalize the study. Some considerations include: **Good:** The range of exposure levels provides sufficient variability in exposure distribution and/or sufficient range or contrasts (e.g., across groups or exposure categories) to detect associations or exposure-response relationships that might be present. The population was exposed to levels expected to have an impact on response. The study population was at risk of developing the outcomes of interest (e.g., ages, lifestage, sex). The timing of outcome ascertainment was appropriate given expected latency for outcome development (i.e., adequate follow-up interval). Evidence was provided of sufficient statistical power (which might include formal power calculations) to observe an effect if it exists. No other concerns raised regarding study sensitivity (e.g., no evidence that results would be attenuated enough to preclude detection of an adverse health effect). **Adequate:** Same considerations as **Good**, except issues are identified that could reduce sensitivity, but they are unlikely to impact the overall findings of the study. **Deficient:** Concerns were raised about the issues described for **Good** that are expected to notably decrease the sensitivity of the study to detect associations for the outcome (i.e., reasonably high likelihood of a false null result). Note: *Deficient* sensitivity indicates that null findings should be interpreted with caution and might not represent a lack of association. **Critically Deficient:** Severe concerns were raised about the sensitivity of the study such that any observed association is uninterpretable (e.g., exposure gradients/contrasts that precluded an ability to distinguish exposure levels between study participants).

[a]The considerations are based on established EPA systematic review methods for evaluation of epidemiology studies (U.S. EPA, 2022b)

6.2.1. Fluoride-Specific Epidemiology Study Evaluation Criteria for the Exposure Measurement Domain

Table 6-2 describes fluoride-specific criteria for evaluation of the exposure measurement domain for studies in the future Fluoride Human Health Toxicity Assessment. Separate criteria are provided for measurements of fluoride in drinking water and urine, other tissues (e.g., teeth, bone, fingernails, hair, blood, and saliva), and for estimated fluoride intake. Multiple exposure measurements from the same publication are evaluated separately and may receive different ratings.

As described in Section 2.4 above, human exposure to fluoride can be estimated in a number of ways that fall into two broad categories – internal measurements of fluoride inside the body, often referred to as biomarkers of exposure (e.g., concentrations of fluoride in urine, blood, or other tissues), and external measurements of fluoride sources outside of the body (e.g., concentrations of fluoride in drinking water). A strong relationship, or correlation, has been reported between concentrations of fluoride in urine and drinking water (Pearson correlation coefficient, $r_p = 0.79$) and between concentrations of fluoride in serum and drinking water ($r_p = 0.62$) across populations (EFSA Scientific Committee, 2025). These strong correlations between measured fluoride biomarkers (urine and serum concentrations) and drinking water provide support that fluoride biomarkers are valid measures of fluoride exposure. Thus, both fluoride concentrations in drinking water and internal biomarker measures of fluoride (e.g., urine concentrations) are considered valid approaches to characterize human fluoride exposure in general, but their validity in each specific study is evaluated, and the identified strengths and limitations are documented.

Several factors that are considered when evaluating the validity of internal and/or external measures of exposure in each study are described in Table 6-2. For example, the analytical method or laboratory technique used in each study is considered since these may vary across laboratories. The analytical method's sensitivity and specificity, as well as the ability of the method to account for all possible sources of fluoride exposure (e.g., through water, diet, or food) is also evaluated. Additionally, the timing of fluoride exposure measurement in relation to the health outcome measurement is evaluated, specifically considering information about the sensitive or critical window for the health outcome of interest. Measurements of fluoride are subject to within-person variability as exposure levels can change over time. Internal fluoride measurements are also likely to differ across individuals based on inter-individual differences in fluoride elimination from the body. The extent to which different approaches of exposure measurement minimize sources of error will also be evaluated. For example, a study measuring water concentrations of fluoride directly captures a relevant fluoride exposure source with a clear interpretation, but one using biomarkers of exposure captures a measure of total fluoride exposure without relying on participant's self-reported water or food intake. Similarly, a study using a 24-hour urine sample provides a more complete picture of a single day's exposure, but a study with multiple spot urine samples may provide more information about typical exposure over time.

Table 6-2. Fluoride-Specific Criteria to Guide Study Evaluation for the Exposure Domain

Core Question: Does the exposure measure reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome?
Criteria
Good: Valid exposure assessment methods used, which represent the etiologically relevant time period for effects on neurodevelopmental outcomes (e.g., exposure during a critical developmental window or exposure preceding the evaluation of the outcome). Exposure misclassification is expected to be minimal. ***For measurements in drinking water and urine:*** Use of appropriate analytical method such as ion selective electrode method or ion chromatography (both preferred), or other standard methods such as National Institute for Occupational Safety and Health (NIOSH) Method 8308 or other governmental standard methods; includes documentation of quality assurance/quality control (QA/QC) such as use of blanks or inter-assay variability, calibration, chemical standards, spiking, recovery rate, limit of detection (LOD, or similar value such as limit of quantification, reporting limit, etc.), LOD, percentage less than LOD, and methods used to account for values below the LOD. Individual-level water use data in combination with drinking water concentrations at relevant locations (e.g., home, school, workplace). Water use data should be specific to the population under study (e.g., consumption patterns obtained via validated or reliable water use surveys such as 24-hour or 7-day diary data on intake and other water use practices collected "prospectively" in real-time or near when activities occurred, or "retrospective" recall questionnaires such as what was ingested in the last 24 hours or a typical week). These estimates should be representative of typical exposure patterns during the etiologically relevant window and/or reflect multiple time points if considerable variation is anticipated during these windows. Use of long-term (24-hour) and/or repeated (e.g., multiple time points) urine samples is preferred to better characterize exposure variability, given the relatively short half-life of fluoride in urine (generally in the 4- to 6-hour range). If adjusting for volume dilution, use of an appropriate approach (e.g., creatinine and/or specific gravity) with rationale provided for the adjustment method (Note: urinary dilution is not required). Exposure is characterized for the relevant time period – for example, maternal urinary measures in the appropriate stage of pregnancy for the outcome(s) of interest when evaluating neurodevelopmental endpoints. Exposure measures are more likely to represent long-term exposure (for example, the study population consists of persons residing in the same area for a long period of time, when water fluoride concentrations were known to be relatively stable). ***For measurements in other tissues (e.g., teeth, bone, fingernails, hair, blood, saliva):*** Fluoride measured in teeth or bone under appropriate sampling conditions. Use of an appropriate analytical method; includes documentation of QA/QC such as use of blanks or inter-assay variability, calibration, chemical standards, spiking, recovery rate. Consideration of exposure context via appropriate adjustment and/or modeling (e.g., PBPK model to estimate target tissue concentrations).

Core Question: Does the exposure measure reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome?

Adequate:
Valid exposure assessment methods used, which represent the etiologically relevant time period of interest.
Exposure misclassification might exist but is not expected to greatly change the effect estimate.
For measurements in drinking water and urine:
Same as **Good**, except for one or more of the following:
The analytical method might not be standard/widely accepted, but a method validation study was conducted prior to sample analysis and is expected to be consistent with sound scientific theory and/or accepted approaches.
Might rely on spot urine samples and/or single urine sample.
Estimates of individual-level water use in combination with drinking water concentrations at relevant locations during etiologically relevant period, but water use data are based on external information about consumption data not specific to the population (e.g., EPA Exposure Factors Handbook (U.S. EPA, 2019a)), or no specific water use information provided.
Water concentration data representative of the etiologically relevant period.
For measurements in other tissues (e.g., teeth, bone, fingernails, hair, blood, saliva):
Fluoride measured in plasma and saliva under appropriate sampling conditions.
Use of an appropriate analytical method; includes documentation of QA/QC such as use of blanks or inter-assay variability, recovery rate.
Consideration of exposure context via appropriate adjustment and/or modeling (e.g., PBPK model to estimate target tissue concentrations).
For estimated intake:
Applied in a population not exposed to fluoride through fluoridated salt.
Use of appropriate instrument to evaluate intake from food, beverage (including drinking water), and dental products over a relevant time period. For example, a validated food frequency questionnaire or duplicate diet sample might be appropriate depending on the study population.
For infants, consideration of breastfeeding, formula intake, and any other potential routes of exposure.
Use of a validated and accurate database of fluoride content for food, beverage, and dental items or use of an appropriate analytic approach to measure fluoride content in samples of food, beverage, and dental items with reporting of QC information.

Deficient:
Valid exposure assessments used, which represent the etiologically relevant time period of interest. Specific knowledge about the exposure and outcome raise concerns about reverse causality, but there is uncertainty whether it is influencing the effect estimate.
Exposed groups are expected to contain a notable proportion of unexposed or minimally exposed individuals, the method did not capture important temporal or spatial variation, or there is other evidence of exposure misclassification that would be expected to notably change the effect estimate.
Studies relying on qualitative groupings of exposure status with some measurement or confirmation of the difference in exposure status (e.g., comparison between high vs. low fluoride drinking water communities with at least some measurement data to indicate the differentiation in fluoride levels between the two communities).
For measurements in drinking water and urine:
Did not identify analytical methods used to measure fluoride in urine or water samples.
Failure to report any QA/QC information (examples of reporting elements in **Good**).
Exposure measures are unlikely to represent exposures in the etiologically relevant period (for example, residential mobility is not addressed in a study population where participants might have moved between residences with varying levels of fluoride in drinking water).
For measurements in other tissues (e.g., teeth, bone, fingernails, hair, blood, saliva):

Core Question: Does the exposure measure reliably distinguish between levels of exposure in a time window considered most relevant for a causal effect with respect to the development of the outcome?
Fluoride measured in hair or nails. Did not identify analytical methods used to measure fluoride in tissues. Failure to report LOD, percentage less than LOD, and methods used to account for values below the LOD. ***For estimated intake:*** Applied in a population exposed to fluoride through fluoridated salt. Did not identify instruments or methods used to evaluate fluoride intake. Lack of consideration of major sources of exposure (e.g., only evaluated diet and drinking water, but not other beverages such as tea). **Critically Deficient:** Exposure measurement does not characterize the etiologically relevant time period of exposure or is not valid. Evidence that reverse causality is very likely to account for the observed association. Exposure measurement was not independent of outcome status. Studies relying entirely on qualitative or subjective exposure measures with no information on reliability for the study population and no confirmation of exposure contrast between groups. ***For measurements in drinking water and urine:*** Analytical methodology is not scientifically appropriate for the chemical and media being analyzed (e.g., method not sufficiently sensitive, not specific to the chemical, out of date). QA/QC issues have been identified which prevent reliable interpretation of the study findings (e.g., notable interference by co-occurring compounds). Studies relying entirely on qualitative exposure measures (e.g., residence in a community with or without community fluoridation) with no confirmation of a contrast in exposure between groups or individuals. ***For measurements in other tissues (e.g., teeth, bone, fingernails, hair, blood, saliva):*** Analytical methodology is not scientifically appropriate for the chemical and media being analyzed (e.g., method not sufficiently sensitive, not specific to the chemical, out of date). QA/QC issues have been identified which prevent reliable interpretation of the study findings (e.g., notable interference by co-occurring compounds). Failure to report correspondence of biomarker with more established methods of fluoride exposure assessment or exposure context. ***For estimated intake:*** Studies relying entirely on qualitative or subjective exposure measures with no information on reliability or suitability for the study population.

6.2.2. Neurodevelopmental Effects Outcome-Specific Epidemiology Study Evaluation Criteria for Outcome Ascertainment and Confounding Domains

The outcome-specific criteria developed for the future Fluoride Human Health Toxicity Assessment for evaluating studies of neurodevelopmental effects are described below. These relate to outcome ascertainment (e.g., timing of outcome evaluation and the instrument(s) used for outcome evaluation) and confounding domains.

6.2.2.1. Outcome Ascertainment

Table 6-3 provides criteria used during study evaluation of the outcome ascertainment domain for studies of neurodevelopmental effects. Although the epidemiological evidence is assessed according to categories of tests, neurodevelopmental outcomes are not discrete, and there is interconnection across categories. For example, attention and response inhibition, which is categorized under tests of attention, impulse control, and externalizing/internalizing behaviors, are core features of both executive function and cognitive function (e.g., scores on intellectual function will be better if an individual is attending to the test).

The studies in the neurodevelopmental database employed a number of different test types and assessment methods. This heterogeneity can pose challenges for comparisons across studies, as each of these tests assesses neuropsychological function somewhat differently. For example, performance-based tests of behavior are more objective measures than behavioral rating scales (often based on parental or teacher reporting). However, tests offer only a snapshot of behavior at a single time point in a single environment and therefore do not reflect behavior over time or across different settings (e.g., home, work). Relatedly, the validity and reliability of the neuropsychological tests varied across studies (NIEHS, 2022). Even when the instrument has been validated for use in a specific population, there could be challenges when applying it in a different group of individuals (e.g., due to language or cultural norms). Consequently, it was important to determine whether tests were normed for the specific study population, including whether the test was developed in the language spoken by the study participants or was translated. Studies also did not consistently report on test conditions, training of examiners in test administration, blinding of examiners to the level of fluoride exposure of participants, interrater reliability, and mode of administration (e.g., computer vs. paper-pencil). Blinding might not be a major concern in some study settings but could present a source of potential bias in situations where participants might have visible manifestations of exposure (e.g., dental fluorosis) or exposure assignment is based in part on residential location. Age at outcome assessment also likely impacted whether an association was detected, particularly in the developing brain, which continues to mature into early adulthood (White et al., 2009).

Table 6-3. Study Evaluation Criteria for Outcome Ascertainment of Neurodevelopmental Effects

Core Question: Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?	
Prompting Questions	**Follow-Up Questions**
Is outcome ascertainment likely to be affected by knowledge of, or presence of, exposure (e.g., consider access to health care, if based on self-reported history of diagnosis)? *For case-control studies:* Is the comparison group without the outcome (e.g., controls in a case-control study) based on objective criteria with little or no likelihood of inclusion of people with the disease? *For diagnosis of disease measures:* Is the diagnosis based on standard clinical criteria? If it is based on self-reporting of the diagnosis, what is the validity of this measure?	Is there a concern that any outcome misclassification is nondifferential, differential, or both? What is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?
Criteria	
Good High certainty in the outcome definition (i.e., specificity and sensitivity), minimal concerns with respect to misclassification. Assessment instrument was validated in a population comparable to the one from which the study group was selected. ***For tests of cognition or intelligence, memory, and learning:*** Use of gold-standard or other validated test instruments (e.g., Wechsler Abbreviated Scale of Intelligence; Bayley Scales of Infant Development). Thorough description of how the test was adapted and validated for tests administered in a different language or culture other than that for which the test was originally developed and standardized (if applicable). Inclusion of outcome measures reflecting specific neuropsychological domains (learning, memory, visual spatial skills, motor function) or subscales of an omnibus test (e.g., Wechsler Intelligence Scale for Children-Fourth Edition IQ subscales: Verbal Comprehension, Perceptual Reasoning, Working Memory, and Processing Speed). ***For tests of younger children, considered factors that might affect the child's state (e.g., sleep, health indicator, time since last fed) or utilized repeated administrations:*** Test system(s) were administered in such a way as to minimize influence of examiner and optimize standardization of administration. Assessments conducted by trained interviewers. Testing shown to be reliable across multiple raters or consistently applied by a single rater (i.e., inter-rater and intra-rater reliability reported), with adjustment for potential differences between raters as appropriate. Adequate description of testing environment (e.g., home, clinic, research site), examiner training, quality control procedures (check related references from the cohort or study group). ***For tests of attention, impulse control, and externalizing/internalizing behaviors:*** Use of gold-standard or other validated test instruments or rating scales (e.g., Bayley Scales of Infant Development; Conners Rating Scale). Where computerized and non-computerized versions of the same instrument are available, prefer use of computerized version. Assessments conducted by trained interviewers (if appropriate).	

Core Question: Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?
Testing shown to be reliable across multiple raters or consistently applied by a single rater (i.e., inter-rater and intra-rater reliability reported), with adjustment for potential differences between raters as appropriate.
Rating scale or performance-based test (preferably both), with inclusion of test subscales or Diagnostic and Statistical Manual of Mental Disorders (DSM) subscales, if available.
Thorough description of how the test was adapted and validated for tests administered in a different language or culture other than that for which the test was originally developed and standardized (if applicable).
For tests of executive function:
Use of gold-standard or other validated test instruments (e.g., Trail Making Test; Wisconsin Card Sorting Test).
Where computerized and non-computerized versions of the same instrument are available, prefer use of computerized version.
Assessments conducted by trained interviewers.
Testing shown to be reliable across multiple raters or consistently applied by a single rater (i.e., inter-rater and intra-rater reliability reported), with adjustment for potential differences between raters as appropriate.
Rating scale or performance-based test (preferably both), with inclusion of test subscales or DSM subscales, if available.
Thorough description of how the test was adapted and validated for tests administered in a different language or culture other than that for which the test was originally developed and standardized (if applicable).
Assessment at age ≥ 10 years.
For tests of social cognition and social behavior (including behaviors related to autism spectrum disorders):
Use of gold standard or other validated test instruments or rating scales (e.g., Conners Comprehensive Behavior Rating Scales).
Where computerized and non-computerized versions of the same instrument are available, prefer use of computerized version.
Assessments conducted by trained interviewers (if appropriate).
Testing shown to be reliable across multiple raters or consistently applied by a single rater (i.e., inter-rater and intra-rater reliability reported), with adjustment for potential differences between raters as appropriate.
For researcher-administered tests, description of the test environment (ideally training of assessment administrators and locations of observations [e.g., home vs. clinical test room]).
Rating scale or performance-based test (preferably both), with inclusion of test subscales.
Thorough description of how the test was adapted and validated for tests administered in a different language or culture other than that for which the test was originally developed and standardized (if applicable).
For tests of sensory and motor function:
Use of gold-standard or other validated test instruments (e.g., Sensory Integration Praxis Test; Beery-Buktenica Developmental Test of Visual-Motor Integration).
Thorough description of how the test was adapted and validated for tests administered in a different language or culture other than that for which the test was originally developed and standardized (if applicable).
Inclusion of outcome measures reflecting specific neuropsychological domains or subscales of an omnibus test.
For tests of younger children, considered factors that might affect the child's state (e.g., sleep, health indicator, time since last fed) or utilized repeated administrations.
Test system(s) were administered in such a way as to minimize influence of examiner and optimize standardization of administration.
Assessments conducted by trained interviewers.
Testing shown to be reliable across multiple raters or consistently applied by a single rater (i.e., inter-rater and intra-rater reliability reported), with adjustment for potential differences between raters as appropriate.

Core Question: Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?

Adequate description of testing environment (e.g., home, clinic, research site), examiner training, quality control procedures (check related references from the cohort or study group).

Adequate

Moderate confidence that outcome definition was specific and sensitive, some uncertainty with respect to misclassification but not expected to greatly change the effect estimate.

Assessment instrument was validated but not necessarily in a population comparable to the study group.

For tests of cognition or intelligence, memory, and learning; for tests of attention, impulse control, and externalizing/internalizing behaviors; for tests of executive function; for tests of social cognition and social behavior (including behaviors related to autism spectrum disorders); for tests of sensory and motor function:

Similar to **Good** but might not have included the same level of detail on test characteristics and administration.

Deficient

Outcome definition was not specific or sensitive.

Uncertainty regarding validity of assessment instrument.

For tests of cognition or intelligence, memory, and learning:

Failure to blind to exposure or to report on blinding to exposure ONLY for parent-completed behavioral rating scales or self-reported outcomes, or in situations where blinding would present substantial difficulty (e.g., children had severe dental fluorosis).

Reliance on self or parental report only.

No description of how the test was adapted and validated for tests administered in a different language or culture.

No adjustment for children's video game usage or for adults' computer experience for computer-administered tests.

For tests of attention, impulse control, and externalizing/internalizing behaviors:

Failure to blind to exposure or to report on blinding to exposure ONLY for parent-completed behavioral rating scales or self-reported outcomes, or in situations where blinding would present substantial difficulty (e.g., children had severe dental fluorosis).

Use of non-validated self-report forms.

No description of how the test was adapted and validated for tests administered in a different language or culture.

No adjustment for children's video game usage or for adults' computer experience for computer-administered tests.

For tests of executive function:

Failure to blind to exposure or to report on blinding to exposure ONLY for parent-completed behavioral rating scales or self-reported outcomes, or in situations where blinding would present substantial difficulty (e.g., children had severe dental fluorosis).

Use of non-validated self-report forms.

No description of how the test was adapted and validated for tests administered in a different language or culture.

No adjustment for children's video game usage or for adults' computer experience for computer-administered tests.

Assessment at <7 years of age.

For tests of social cognition and social behavior (including behaviors related to autism spectrum disorders):

Failure to blind to exposure or to report on blinding to exposure ONLY for parent-completed behavioral rating scales or self-reported outcomes, or in situations where blinding would present substantial difficulty (e.g., children had severe dental fluorosis).

Reliance on self-report.

No description of how the test was adapted and validated for tests administered in a different language or culture.

No adjustment for children's video game usage or for adults' computer experience for computer-administered tests.

Core Question: Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?

For tests of sensory and motor function:

Failure to blind to exposure or to report on blinding to exposure ONLY for parent-completed behavioral rating scales or self-reported outcomes, or in situations where blinding would present substantial difficulty (e.g., children had severe dental fluorosis).

Reliance on self- or parental report only.

No description of how the test was adapted and validated for tests administered in a different language or culture.

No adjustment for children's video game usage or for adults' computer experience for computer-administered tests.

Critically Deficient

Invalid/insensitive marker of outcome.

Outcome ascertainment is very likely to be affected by knowledge of, or presence of, exposure.

Note: Lack of blinding should not be automatically construed to be ***critically deficient.***

For tests of cognition or intelligence, memory, and learning; for tests of attention, impulse control, and externalizing/internalizing behaviors; for tests of executive function; for tests of social cognition and social behavior (including behaviors related to autism spectrum disorders); for tests of sensory and motor function:

Use of test instruments established to be invalid or unreliable.

Administration of a test to participants outside the age range for which the test has been validated.

6.2.2.2. Potential Confounding

Table 6-4 provides criteria for evaluation of the confounding domain for studies of neurodevelopmental effects. Study design features and statistical approaches such as multivariable adjustment are used to minimize potential for confounding. Specific adjustment sets that should be included in the analyses depend on the age range of the study population and the exposure context (e.g., general population, relevant co-exposures).

Below are examples of potential confounders (variables associated with both the exposure and outcome; adjustment reduces bias) and risk factors (variables likely only associated with the outcome, indicated with an asterisk* below; adjustment improves precision) that may be considered across different contexts. Judgments about the relevance of each of these variables are made on a study-by-study basis by the evaluator based on available information.

- Age
  - (Green et al., 2020b; Thornton-Evans et al., 2019; Zohoori et al., 2013; Villa et al., 2010)
- Indicator of parental cognition (e.g., maternal and/or paternal IQ or education)
  - (Thomas et al., 2016; Bartels et al., 2002)
- Indicator of socioeconomic status of household/caregiver(s) (e.g., caregiver education, caregiver occupation, household income)
  - (Chi et al., 2024; Turkheimer et al., 2003)
- Race/ethnicity
  - (Hefferon et al., 2024; Krieg et al., 2001; Kramer et al., 1995)
- Breastfeeding/duration
  - (Plunkett et al., 2021; Zohoori et al., 2019; Levy et al., 1995; Ekstrand et al., 1981)
- Physical activity[5]
  - (Erickson et al., 2019)
- Co-occurring exposures via food or drinking water, such as:
  - Iodine[6]
    - (Qian et al., 2005)
  - Lead (Pb)[7]
    - (U.S. EPA, 2024d)
  - *Methylmercury
    - (Axelrad et al., 2007; Trasande et al., 2005)
  - Arsenic[8]

[5]Elevated physical activity is presumed to be associated with elevated fluoride exposure via increased water intake.

[6]Consideration of iodine may be important in (a) populations at risk of iodine insufficiency and (b) areas such as Mexico and Denmark (lists of relevant countries can be found in EFSA Scientific Committee (2025) and Pollick (2013) where fluoridated salt may present a co-occurring exposure).

[7]Lead (Pb) may be associated with fluoride exposure among populations drinking tap water that is fluoridated and delivered via Pb pipes.

[8]Inorganic arsenic is more likely to be an issue in certain geographic regions including China, India and Mexico but may sometimes be inferred from groundwater quality maps.

    - (U.S. EPA, 2025c; Limón-Pacheco et al., 2018; Peterson et al., 2016; González-Horta et al., 2015)
- *Sex
    - (Green et al., 2020a)
- *Indicator of home/caregiving environment (e.g., HOME score)
    - (Ronfani et al., 2015; Bradley and Caldwell, 1980)
- *Birth-related factors (e.g., birth weight, preterm birth)
    - (Kerr-Wilson et al., 2012; Shenkin et al., 2004)
- *Maternal use of alcohol/tobacco/illicit drugs during pregnancy
    - (Corrêa et al., 2021; Morrow et al., 2006; Streissguth et al., 1990; Streissguth et al., 1989)

Table 6-4. Study Evaluation Criteria for Confounding of Neurodevelopmental Effects

Core Question: Is confounding of the effect of the exposure likely?	
Prompting Questions	**Follow-Up Question**
Is confounding adequately addressed by considerations in: Participant selection (matching or restriction)? Accurate information on potential confounders and statistical adjustment procedures? Lack of association between confounder and outcome, or confounder and exposure in the study? Information from other sources? Is the assessment of confounders based on a thoughtful review of published literature, potential relationships (e.g., as can be gained through directed acyclic graphing), and minimizing potential overcontrol (e.g., inclusion of a variable on the pathway between exposure and outcome)?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?
Criteria	
Good Conveys strategy for identifying potential confounders. This might include: a priori biological considerations, published literature, causal diagrams, or statistical analyses; with recognition that not all "risk factors" are confounders. All key confounders are included (or excluded based on analyses documenting that they are not relevant to the population under study). Potentially relevant confounders are evaluated and included appropriately if demonstrated to be sources of substantial confounding. This often will include: Presenting the distribution of potential confounders by levels of the exposure of interest and/or the outcomes of interest (with amount of missing data noted). Consideration that potential confounders were rare among the study population, or were expected to be poorly correlated with exposure of interest. Consideration of the most relevant functional forms of potential confounders. Examination of the potential impact of measurement error or missing data on confounder adjustment. Presenting a progression of model results with adjustments for different potential confounders, if warranted. Inclusion of potential confounders in statistical models not based solely on statistical significance criteria (e.g., $p < 0.05$ from stepwise regression). Does not include variables in the models that are likely to be influential colliders or intermediates on the causal pathway. **Adequate** Similar to **Good,** but might not have included some relevant key confounders, or less detail might be available on the evaluation of confounders (e.g., sub-bullets in **Good**). It is possible that residual confounding could explain part of the observed effect, but concern is minimal. **Deficient** Some key confounders included but other confounders (or a particular confounder specifically relevant to this study population) not adjusted or accounted for. Assessed an outcome based on report of medical diagnosis that would have required access to a health professional (e.g., autism, attention deficit hyperactivity disorder (ADHD), depression) and failed to consider some marker of socioeconomic status (e.g., maternal education, household income, marital status, crowding, poverty, job status) as a potential confounder. Does not include variables in the models that are likely to be influential colliders or intermediates on the causal pathway. Any of the following:	

Core Question: Is confounding of the effect of the exposure likely?
The potential for bias to explain some of the results is high based on an inability to rule out residual confounding, such as a lack of demonstration that key confounders of the exposure-outcome relationships were considered. Descriptive information on confounders (e.g., their relationship relative to the outcomes and exposure levels) are not presented. Strategy of evaluating confounding is unclear or is not recommended (e.g., only based on statistical significance criteria or stepwise regression [forward or backward elimination]). **Critically Deficient** Includes variables in the models that are colliders and/or intermediates in the causal pathway, indicating that substantial bias is likely from this adjustment Substantial confounding is likely present and not accounted for, such that all of the results were most likely due to bias (e.g., no consideration of confounders judged to be essential in the study population).

6.2.3. Dental Effects Outcome-Specific Epidemiology Study Evaluation Criteria for Exposure Measurement, Outcome Ascertainment, Confounding, and Analysis Domains

Outcome-specific criteria for evaluating studies of dental fluorosis are described in the sections below. Modifications to the standard evaluation criteria were made primarily with respect to the outcome ascertainment domain; however, there are also criteria for the exposure measurement, confounding, and analysis domains that are specific to the evaluation of dental fluorosis studies.

6.2.3.1. Outcome Ascertainment

For the outcome ascertainment domain, information about the conditions of dental examinations as well as inter- or intra-examiner reliability are important in assessing confidence in outcome ascertainment during study evaluation. A primary focus of this domain is the test instruments used to assess the presence and severity of dental fluorosis in study populations. Validated instruments include the Dean's Index, the TFI, and the TSIF (U.S. EPA, 2010a; NRC, 2006). Because some dental fluorosis indices do not provide clear individual-level definitions of severity, it is important that studies provided definitions for fluorosis cases and severity in the assessed study populations to ensure consistency of interpretation. Differences between studies with respect to the definition of dental fluorosis and its severity are also important to consider because this can lead to different incidence rates for dental fluorosis in the population. For consistency, in the future Fluoride Human Health Toxicity Assessment, moderate fluorosis is considered to correspond to a score of 3 on the Dean's Index and a score of 3 to 4 on either the TFI or TSIF. Severe fluorosis is defined as a score of 4 on the Dean's Index, 5 to 9 on the TFI, and 5 to 7 on the TSIF (U.S. EPA, 2010a; NRC, 2006). These definitions are consistent with those used in the previous, peer-reviewed fluoride assessment by the EPA (U.S. EPA, 2010a). This standardization of severity measures enables study results with different dental fluorosis indices to be directly compared. Additional context for the evaluation of dental fluorosis is provided in the following paragraphs and Table 6-5 describes specific criteria that apply to the evaluation of dental fluorosis studies.

There are several validated indices of fluorosis commonly used in the literature, including Dean's Index, the TFI, TSIF, and the Fluorosis Risk Index (FRI). These indices record observations at the person- (e.g., Dean's Index), tooth- (e.g., TF index), or tooth surface-level (e.g., TSIF and FRI). Tooth- and tooth surface-level measurements can provide greater sensitivity and allow more refined investigations of periods of susceptibility. However, they have less inter- and intra-evaluator reliability and may be subject to greater misclassification due to the larger number of measurements made on each study participant. Yet, when considering severe or moderately severe fluorosis, there is a high degree of consistency across fluorosis indices. Notably, although the Community Fluorosis Index is calculated based on individual-level Dean's Index measurements, its validity has been questioned because of "the statistical basis for using the arithmetic mean to calculate… [it when] the classification is based on an ordinal and not an interval scale" (Clarkson, 1989).

Timing of outcome ascertainment is considered when evaluating studies on fluorosis. Both deciduous (i.e., baby) and permanent teeth can develop fluorosis; however, the future Fluoride Human Health Toxicity Assessment is focused on fluorosis in permanent teeth. Permanent teeth begin to erupt at approximately age 6–7 and this continues until approximately age 13. In general, the first molar and incisors erupt first, followed by the canines, premolars, and second molars. An exception to this timeline is the third molars (i.e., wisdom teeth), which erupt around age 17–21.

There is evidence that tooth fluorosis lesions (especially mild or moderate lesions) fade in severity into adulthood. Therefore, determinations of fluorosis prevalence are best made from late childhood (e.g., 8–11 years) to adolescence (e.g., >11 years) with numbers of deciduous, decayed, or missing teeth noted. Notably, because late-erupting teeth are also more susceptible to developing fluorosis, measurements in populations of adolescents with all or most of the adult teeth (e.g., 12–14 years) may be preferred.

Table 6-5. Evaluation Criteria for Outcome Ascertainment for Dental Effects

Evaluation Considerations for Outcome Ascertainment	
Core Question: Does the outcome measure reliably distinguish the presence or absence (or degree of severity) of the outcome?	
Prompting Questions	**Follow-Up Questions**
Is outcome ascertainment likely to be affected by knowledge of, or presence of, exposure (e.g., consider access to health care, if based on self-reported history of diagnosis)? ***For case-control studies:*** Is the comparison group without the outcome (e.g., controls in a case-control study) based on objective criteria with little or no likelihood of inclusion of people with the disease? ***For mortality measures:*** How well does cause of death data reflect occurrence of the disease in an individual? How well do mortality data reflect incidence of the disease? ***For diagnosis of disease measures:*** Is the diagnosis based on standard clinical criteria? If it is based on self-report of the diagnosis, what is the validity of this measure? ***For laboratory-based measures (e.g., hormone levels):*** Is a standard assay used? Does the assay have an acceptable level of inter-assay variability? Is the sensitivity of the assay appropriate for the outcome measure in this study population?	Is there a concern that any outcome misclassification is nondifferential, differential, or both? What is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?
Criteria	
Good High certainty in the outcome definition (i.e., specificity and sensitivity), minimal concerns with respect to misclassification. Assessment instrument was validated in a population comparable to the one from which the study group was selected. ***For measurement of dental fluorosis:*** Use of validated test instruments (e.g., Dean's Index, the Thylstrup-Fejerskov (TF) index, Tooth Surface Index of Fluorosis (TSIF), the Fluorosis Risk Index (FRI)). Case and severity definitions clearly stated (e.g., fluorosis status determined by the two most severely affected teeth on Dean's Index). Examiner reliability is reported (e.g., inter and intra-reviewer agreement, kappa statistics) and, if relevant, conditions of field examination reported (e.g., use of specific lighting, drying of the teeth, reference images). Examinations are performed on permanent teeth in children or adolescents.	
Adequate *Moderate* confidence that outcome definition was specific and sensitive, some uncertainty with respect to misclassification but not expected to greatly change the effect estimate. Assessment instrument was validated but not necessarily in a population comparable to the study group. ***For measurement of dental fluorosis:*** Similar to **Good** but may not have included the same level of detail on test characteristics and administration.	
Deficient	

Evaluation Considerations for Outcome Ascertainment
Outcome definition was not specific or sensitive. Uncertainty regarding validity of assessment instrument. ***For measurement of dental fluorosis:*** Outcome definition was not specific or sensitive. For example, outcome is reported as 'no fluorosis' vs. 'any fluorosis' without further information on distribution of severity or extent. Case and severity definitions are not clearly defined or examinations are not limited to children or adolescents. Uncertainty regarding validity of assessment instrument (e.g., Community Fluorosis Index (CFI)). Outcome assessment includes deciduous (baby) teeth.
Critically Deficient Invalid/insensitive marker of outcome. Outcome ascertainment is very likely to be affected by knowledge of, or presence of, exposure. ***For measurement of dental fluorosis:*** Use of test instruments established to be invalid or unreliable. Outcome ascertainment is very likely to be affected by knowledge of, or presence of, exposure. Note: Lack of blinding should not be automatically construed to be *Critically Deficient*.

6.2.3.2. Exposure Measurement

Fluoride-specific exposure criteria are reported above in Section 6.2.1 and applied to the evaluation of dental fluorosis studies. For more context regarding relevant exposure periods for the development of dental fluorosis, evaluators are provided with the information and citations in the following paragraph.

For dental fluorosis, the etiologically relevant period of exposure is during the maturation of the enamel. In general, for permanent (adult) teeth, the relevant exposure period for fluorosis development appears to be from infancy through approximately age 8, with some evidence that exposures occurring after this age are not strongly linked to fluorosis incidence (e.g., (NRC, 2006; DenBesten and Thariani, 1992; Ishii and Suckling, 1991)). For studies that focus on specific teeth or zones of the mouth, more refined exposure categories may be relevant. For example, early-erupting teeth (central incisors and first molars) appear to be at highest risk of developing fluorosis after exposures occurring during the first 2–3 years of life. Later-erupting teeth (canines, premolars, second molars) appear to be at highest risk between the period of 2–8 years, though this may vary slightly depending on the specific tooth and tooth surface. Because of the longer etiologically relevant period of exposure for later-erupting teeth (and typically higher doses of fluoride in older children), these later-erupting teeth may be more sensitive to developing fluorosis compared with earlier erupting teeth. Despite these differences, the NRC (2006) report concludes that "because the severity of fluorosis is related to the duration, timing, and dose of fluoride intake, cumulative exposure during the entire maturation stage, not merely during critical periods of certain types of tooth development, is probably the most important exposure measure to consider when assessing the risk of fluorosis" (DenBesten, 1999).

In addition, some studies in this literature database use older methods of chemical analysis that are susceptible to chemical interferences, namely, colorimetric methods using zirconium-alizarin. See resources in Appendix D for more information.

6.2.3.3. Confounding

Fluoride exposure is the only cause of dental fluorosis. Therefore, the number of confounding variables to account for is limited. Potential confounders to include in analyses depend on the age range of the study population and the study context (e.g., population factors, exposure scenario, or statistical analysis). Listed below are examples of potential confounders or risk factors of the exposure and/or outcome that might be considered across different studies. Because many of these variables are risk factors of either the exposure or the outcome, their potential role as a confounder is determined on a study-by study basis by the evaluator based on available information.

Potential confounders and risk factors include:

- Age (CDC, 2019; Mascarenhas, 2000). Fluoride intake varies across age groups due to changes in feeding practices (e.g., breast or formula feeding, changing water consumption and dietary patterns). Age also relates to the risk of dental fluorosis because fluorosis is measured on individual teeth (or surfaces) and having more teeth to observe increases the probability of identifying fluorosis. In addition, later-erupting teeth appear more susceptible to developing dental fluorosis than early-erupting teeth. Note: many

studies will restrict and/or match participants based on age and evaluators should consider whether additional adjustment is needed.

- For studies of water fluoride, other sources of fluoride may be relevant, such as (U.S. EPA, 2010a; Mascarenhas, 2000):
    - Breastfeeding/formula feeding/infant feeding practices
    - Toothbrushing with fluoridated toothpaste, use of fluoridated supplements/tablets, use of fluoridated mouthwash in early childhood (e.g., <3 years) or other fluoride treatments
    - Fluoridated salt
    - Tea consumption
    - Excessive dietary fluoride or fluoride in bottled water or sodas
- For studies of water fluoride, factors related to water intake may be relevant, such as (U.S. EPA, 2011a, 2010a; Mascarenhas, 2000):
    - Sex
    - Climate or elevation of the study population
    - BMI
    - Race and ethnicity (or nativity) (CDC, 2019; Mascarenhas, 2000)
    - Household/parental socioeconomic status variables (e.g., parental income, education, insurance status) (CDC, 2019; Mascarenhas, 2000)
    - Use of tetracycline antibiotics or amoxicillin (Hong et al., 2011)
    - Malnutrition (U.S. EPA, 2010a)

Note: In some studies where cohorts may span a wide range of birth years, accounting for time could capture differences in early life exposures due to changing recommendations around water fluoridation, availability of dental treatments/products, or fluoride supplement content and usage.

6.2.3.4. Analysis

Table 6-6 describes evaluation criteria for the analysis of dental fluorosis studies. Information relevant to evaluation of analysis includes, but is not limited to, the extent (and if applicable, treatment) of missing data for exposure, outcome, and confounders, approach to modeling, classification of exposure and outcome variables (continuous vs. categorical), testing of assumptions, sample size for specific analyses, and relevant sensitivity analyses. Information specific to the characterization of dental fluorosis and the evaluation of descriptive studies that present tabular results are provided for evaluators. Note that studies may be given a rating of "Not Applicable," which describes studies that do not perform statistical analyses but do report data in a manner that can be used for some dose-response analyses. This rating is not considered a downgrade.

Table 6-6. Study Evaluation Criteria for Analysis of Dental Studies

Core Question: Does the analysis strategy and presentation convey the necessary familiarity with the data and assumptions?	
Prompting Questions	**Follow-Up Question**
Are missing outcome, exposure, and confounder data recognized, and if necessary, accounted for in the analysis? Does the analysis appropriately consider variable distributions and modeling assumptions? Does the analysis appropriately consider subgroups of interest (e.g., based on variability in exposure level or duration or susceptibility)? Is an appropriate analysis used for the study design? Is effect modification considered, based on considerations developed a priori? Does the study include additional analyses addressing potential biases or limitations (i.e., sensitivity analyses)?	If there is a concern about the potential for bias, what is the predicted direction or distortion of the bias on the effect estimate (if there is enough information)?
Criteria	
Good Use of an optimal characterization of the outcome variable (analysis includes disease severity, where possible). Quantitative results presented (effect estimates and confidence limits or variability in estimates (e.g., standard error, standard deviation); i.e., not presented only as a p-value or "significant"/"not significant"). Descriptive information about outcome and exposure provided (where applicable). Amount of missing data noted and addressed appropriately (discussion of selection issues—missing at random vs. differential). Where applicable, for exposure, includes LOD (and percentage below the LOD), and decision to use log transformation. Includes analyses that address robustness of findings, e.g., examination of exposure-response (explicit consideration of nonlinear possibilities, quadratic, spline, or threshold/ceiling effects included, when feasible); relevant sensitivity analyses; effect modification examined based only on a priori rationale with sufficient numbers. No deficiencies in analysis evident. Discussion of some details may be absent (e.g., examination of outliers).	
Adequate Similar to **Good**, except: Descriptive information about exposure provided (where applicable) but may be incomplete; might not have discussed missing data, cutpoints, or shape of distribution. Includes analyses that address robustness of findings (examples in good), but some important analyses are not performed.	
Deficient Does not conduct analysis using optimal characterization of the outcome variable (e.g., the analysis combines any fluorosis instead of distinguishing between mild, moderate, and/or severe fluorosis). Descriptive information about exposure levels, including variability. not provided (where applicable). Effect estimate and p-value presented, without standard error or confidence interval. Results presented as statistically "significant"/"not significant."	
Critically Deficient Results of analyses of effect modification examined without clear a priori rationale and without providing main/principal effects (e.g., presentation only of statistically significant interactions that were not hypothesis driven). Analysis methods are not appropriate for design or data of the study.	
Not Applicable: No statistical analysis is presented, but tabular data report numbers of participants with dental fluorosis by exposure.	

6.3. Pharmacokinetic (PK) Model Evaluation

6.3.1. Background

PK modeling uses mathematical equations to describe how a chemical is absorbed into the body, how it moves throughout the body, and how it is eliminated from the body, including its ultimate excretion (Campbell et al., 2012). Metabolism is typically part of the elimination process for many chemicals but not for fluoride (Institute of Medicine, 1997b). Therefore, metabolism will not be considered for evaluations conducted in the future Fluoride Human Health Toxicity Assessment.

PK models simplify anatomical complexity by turning the body into one compartment or a series of compartments (Brochot et al., 2005). Within each compartment, the quantity of chemical present at any given moment is determined by the rates of transfer into and out of the compartment over time. In its simplest form, a PK model treats the entire body as a single compartment, where it specifically describes the concentration of the chemical in whole blood or blood plasma and the excretion rate at any moment remains proportional to the concentration in the blood. PK model complexity can increase through inclusion of additional compartments representing specific organs or tissues, or groups of organs or tissues, in the body (Nestorov et al., 1998).

A PBPK model uses physiologically relevant organ volumes or masses and blood flow rates, typically obtained from anatomical and physiological data that are independent of the chemical being modeled, to define transport of a chemical throughout the body (U.S. EPA, 2006). Additionally, PBPK model compartments are linked together in an anatomical fashion to describe how a chemical moves throughout the body through mathematical descriptions of absorption compartments (e.g., stomach for ingestion, lungs for inhalation) and excretion compartments (e.g., kidneys for urinary excretion, feces for biliary excretion). Finally, because the blood flow rates, organ volumes, and urinary excretion rates used in PBPK models are based on physiological data that have been collected from a wide variety of individuals, lifestage- and population-specific values can be used to predict body concentrations for sensitive subpopulations such as children or those with impaired renal function.

A common assumption in PBPK models is that the chemical is evenly distributed in each organ or tissue (group) (U.S. EPA, 2006). For example, the chemical concentration is assumed to be the same in different parts of the liver or in the bones of different parts of the body. However, when chemical-specific PK data or consideration of biological mechanisms indicates that such an assumption will not accurately account for a chemical's distribution in the body, tissue compartments may be sub-divided, such as between the bone surface, where the chemical might rapidly exchange with the circulating blood, and the underlying bone with which the exchange might be much slower.

In a dose-response analysis, PBPK models allow for conversion between an external dose (e.g., the milligrams of the chemical ingested per day) and the internal concentrations in blood or other tissues or in excreta, such as urine, resulting from that external dose. The conversion between external dose and internal concentration can happen in the "forward" manner, where a chemical's external dose (amount, frequency, and route of exposure) is used as an input and

converted with the model into an internal dose (concentration in blood or urine over time). Conversely, these models can also operate in the "reverse" direction, where an external dose (e.g., amount ingested) is estimated from a biomarker of exposure (e.g., a concentration measured in blood or urine). Together, these PK/PBPK approaches are used to derive human-relevant dose metrics, associated with some degree or lack of toxicity, from which the RfD can be derived. PK/PBPK models are the preferred approach for relating external and internal measures of exposure, and they enable RfD derivation based on internal dose (U.S. EPA, 2006). PK models are also a preferred approach for evaluating variability and sensitivity within a population (U.S. EPA, 2014b), such as lifestage-specific changes in ADME, and are routinely used in the EPA toxicity assessments (U.S. EPA, 2024a, b)

6.3.2. Fluoride-Specific PK Models

The 2010 EPA fluoride assessment (U.S. EPA, 2010a) derived points of departure based only on fluoride concentrations in drinking water; PK modeling was not used to relate observed biomarker concentrations (such as fluoride in blood or urine) with externally applied doses (total amounts of fluoride ingested). EPA plans to use PK calculations in the future Fluoride Human Health Toxicity Assessment to relate observed biomarkers concentrations to the total amount ingested. Dose-response analysis using biomarker concentrations requires a PK analysis to convert measured concentrations in these biomatrices to a total amount of fluoride exposure.

6.3.2.1. Model Considerations

When performing dose-response analysis, a PK or PBPK model should be used when an applicable one exists and no equal or better alternative for dosimetric extrapolation is available (U.S. EPA, 2011b). Any models used should accurately reflect current scientific knowledge and be translated into reproducible, transparent code or spreadsheet equations using a computational language familiar to experts. An existing PBPK model might be employed or adapted or a new PBPK model or an alternate quantitative approach can be developed. For example, an existing model might be modified to include a specific organ or tissue, or to simulate different routes of exposure.

Any PK/PBPK model used for the future Fluoride Human Health Toxicity Assessment should have a defensible biological basis or model structure (U.S. EPA, 2006), show agreement with in vivo human PK data, and should match the toxicity assessment's needs. The EPA will first identify existing fluoride PK/PBPK models in the literature and evaluate those models for use in the toxicity assessment. Preferably, a suitable fluoride pharmacokinetic approach for the future toxicity assessment will:

1. Reflect lifestage-dependent bone uptake and turnover which can substantially impact internal and urinary concentrations,
2. Capture lifestage differences in urinary excretion since renal clearance varies from infancy to adulthood (including pregnancy), and
3. Account for maternal-fetal transport (and if necessary, lactational transfer) to characterize exposure during pregnancy and early life.

While PBPK modeling is the preferred approach for deriving human equivalent oral exposures according to the hierarchy of approaches outlined in EPA guidelines (U.S. EPA, 2011b), useful information might be derived from simpler model types that focus on targeted organ systems in the body or empirical analysis of PK data by algebraic curve-fitting.

The initial selection of the PK model or analytic approach will depend on the degree of biological representation, the extent to which the model is fit-for-purpose, and the ability to adapt the model or approach to the needs of the future Fluoride Human Health Toxicity Assessment. For example, a PBPK model that directly accounts for urinary excretion by the kidney can incorporate lifestage-specific changes in urinary flow, whereas classical PK models depend on fitting to observed data. An existing PK/PBPK model that cannot be replicated based on published information (e.g., model details are not available) or uses proprietary code does not meet the requirements for reproducibility and transparency and would not be considered further. Candidate PK/PBPK models can be first identified based on their availability, the suitability or adaptability of their model structure, and the extent to which they have already been demonstrated to predict fluoride PK in humans. Ultimately, the ability to quantitatively link external exposure with internal concentrations (e.g., in blood or bone) or biomarker concentrations (e.g., in urine) will be evaluated for any candidate model or analysis to inform the toxicity assessment's PK approach, and the selected model code or analysis will be made publicly available.

6.3.2.2. Model Evaluation

Following the literature search, the studies categorized as PK/PBPK models will be evaluated in an initial scoping according to the model considerations in Section 6.3.2.1 to identify a subset of models that appear to be best fit-for-purpose. A more in-depth model evaluation will then be implemented to identify the best candidate PK/PBPK models based on availability, suitability, and demonstrated consistency with in vivo human PK data (Table 6-4). A final in-depth evaluation of the PK/PBPK models will be performed to identify those potentially suitable for deriving toxicity values for the toxicity assessment.

The initial scoping process will provide a rapid assessment of the availability, structure, and potential uses of the PBPK/PK models. This initial scoping will not be a full evaluation, but it will serve as a relevancy check. As described previously, PK modeling can range from a full physiological description of the body (PBPK) to simpler, one-compartment PK models. Appendix C.4 provides a review of additional model types that utilize tissue-specific modeling relevant to the future Fluoride Human Health Toxicity Assessment. The literature survey identified PK/PBPK models (Section 4.4) will be summarized and compared against the criteria presented in Section 6.3.2.1 to determine if further evaluation is warranted.

Once the available PK/PBPK models are summarized, the relevant models will be evaluated according to criteria outlined in Appendix F and informed by the ADME literature search (see Appendix C.3). Judgments on the suitability of a model will be made according to scientific and technical criteria. The scientific criteria focus on whether the biology, chemistry, and other information available for chemical pharmacokinetic processes and, to the extent they are included in the model, modes of action are justified (i.e., preferably with citations to support use)

and represented in the model structure and equations. Compartmental PK models are based on the most basic scientific information and principles, so their quality will primarily be judged by the degree to which they match available data. For PBPK models, they will be evaluated against the scientific criteria based on information presented in the model's publication or report along with goodness-of-fit to in vivo data. Because PK/PBPK models aim to quantify fluoride ADME, these scientific criteria will be established following completion of the ADME systematic literature review (see Appendix C.3). Additionally, PK/PBPK models will be evaluated against biomonitoring measurements from real-world exposures in individuals and populations across the lifestages relevant for the toxicity assessment. Technical criteria for model evaluation include availability of the computer programming code and completeness of parameter listing and documentation, including sources for model parameters.

Finally, top candidate models that meet the general scientific and technical criteria will be subjected to an in-depth model code evaluation before application, which will include a thorough review and testing of any computational code obtained from the model authors, along with the scientific assumptions made for the lifestage and population of interest. Alternately, model testing might focus on the EPA's ability to replicate published model results using code previously developed or vetted for this purpose. The in-depth technical and scientific analyses will focus on the (1) clarity in the documentation of model purpose, structure, and biological characterization; (2) validation of mathematical descriptions, parameter values, and computer implementation; and (3) evaluation of each plausible dose metric. The in-depth analysis will be used to evaluate the potential value and cost of developing a new model or substantially revising an existing one. From this evaluation, the EPA will decide on the best path forward for pharmacokinetic modeling approaches to support the toxicity assessment. Any PK model developed by the EPA or any revisions made to an existing model during development of the toxicity assessment will be informed by the extensive ADME review the EPA is undertaking (Appendix C.3) and will be peer reviewed.

7. Data Extraction of Health Effects Studies

Section Overview: Data extraction is the process that the EPA uses to document important details about study results, including how results were collected and analyzed. These details include study design/type, exposure levels, and numerical results that describe the relationship between fluoride exposure and health outcomes. This information is then organized to facilitate comparisons of findings across studies and study populations. This organization helps the EPA determine which studies are best suited to determine the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (i.e., the reference dose or RfD).

Data extraction and management are accomplished using HAWC (Section 6.1). Data elements that might be extracted from epidemiological studies that are prioritized for study evaluation are listed in Table 7-1, though this list could be refined after an initial data extraction workflow pilot phase. Data extraction will focus on studies identified as potential candidates for dose-response analysis after study evaluation (e.g., Table 8-1). Studies deemed to be *Uninformative* during study evaluation are not considered further and, therefore, do not undergo data extraction. *Low*-confidence studies might not undergo data extraction if sufficient data from *Medium*- and *High*-confidence publications are available. Studies also might not undergo extraction if key study information is missing and no response is provided from authors after email outreach.

All findings are considered for extraction regardless of statistical significance to avoid drawing conclusions based on incomplete evidence. However, the extent of the extraction might vary by outcome (i.e., ranging from a short narrative extraction to a full extraction of dose-response and effect size information), depending on the needs for dose-response analysis. For studies on the same or overlapping study populations, one publication is designated as the primary source for data extraction, and the others are considered as secondary publications for each relevant outcome. Factors considered when selecting the primary source include follow-up time, case numbers, and/or publication date. Similarly, decisions about data extraction for *Low*-confidence publications are made based on consideration of the utility and extent of the available evidence.

Table 7-1. Data Elements That Might Be Extracted from Epidemiological Studies

Category	Potential Data Extraction Elements
	Human
Subjects	Study population name/description
	Time frame of study
	Geography (e.g., country, region, state)
	Demographics (e.g., sex, age, or lifestage at exposure, and at outcome assessment)
	Number of subjects
Methods	Study design (e.g., prospective or retrospective cohort, nested case-control study, cross-sectional, population-based case-control study, intervention, case report)
	Health outcome category (e.g., cardiovascular)
	Health outcome (e.g., blood pressure)
	Confounders or modifying factors considered in analysis

Category	Potential Data Extraction Elements
	Chemical name
	Exposure assessment (e.g., blood, urine, hair, air, drinking water, job classification, residence, administered treatment in controlled study)
Results	Exposure levels (e.g., mean, median, measures of variance as presented in paper, such as standard deviation, standard error of the mean, 75th/90th/95th percentile, minimum/maximum)
	Statistical findings (e.g., adjusted β, standardized mean difference, adjusted odds ratio, standardized mortality ratio, relative risk)

Extracted data will be presented in the toxicity assessment and made available for download from HAWC in Excel format when the draft toxicity assessment is publicly released. One team member performs initial data extraction, and a second team member performs a quality assurance/quality control (QA/QC) review of the extraction. Discrepancies in data extraction are resolved by consulting with a third member of the team. Once the extracted data have been verified, they are "locked" to prevent inadvertent changes. Digital rulers, such as WebPlotDigitizer, might be used to extract numerical information from figures if tabular or numerical data are unable to be obtained from study authors. Any use of digital rulers is documented during data extraction.

In addition to extracting quantitative outcome results in their original units from each study, results are transformed, when possible, to a common metric and scale to help compare distinct but related outcomes measured differently across different studies. These standardized effect size estimates facilitate systematic evaluation of whether meta-analysis will be feasible for the toxicity assessment (Section 8.1). Based on metrics across the available studies, a common metric might be used, and the calculation is presented in the toxicity assessment.

For epidemiology studies, the typical approach is to extract adjusted statistical estimates, when possible, rather than unadjusted or raw estimates.

As described in Section 6, the EPA might ask study authors for missing or additional information that would be required to conduct a meta-analysis or dose-response analysis (e.g., missing group size or variance descriptors such as standard deviation or confidence interval). Outreach to study authors is documented and considered unsuccessful if authors do not respond to email or phone outreach within one month of the attempt to contact.

8. Dose-Response Analysis: Toxicity Value Derivation

Section Overview: This section describes how the EPA will perform dose-response analysis for fluoride. Dose-response analysis (also known as exposure-response analysis) is central to how the EPA will determine a toxicity value (i.e., an RfD) for fluoride. The RfD is the amount of fluoride that a person can ingest daily over a lifetime and be unlikely to experience harmful health effects.

During dose-response analysis, the EPA identifies the most reliable and informative human studies that contain usable fluoride exposure and response data, based on study evaluation (described in Section 6) (Section 8.1). These studies are then used to analyze how increasing dose (exposure) impacts health effects (response) (Section 8.2). From this relationship, the EPA identifies a level of fluoride exposure below which harmful effects are unlikely or minimal (the "point of departure" or POD), which serves as a starting point for the derivation of an RfD (Section 8.2.1).

The EPA then applies UFs to the POD to account for five different types of uncertainty and/or variability (e.g., differences among people or exposure durations; gaps in available data) (Section 8.2.2). The EPA considers each of the five types of uncertainty and variability separately and uses peer-reviewed human health risk assessment methodologies and the best-available science to determine an appropriate UF for each. A total or "composite" UF is calculated by multiplying the five individual UFs together. To arrive at an RfD, the POD is divided by the composite UF.

If there are multiple health effects (e.g., both neurodevelopmental effects and dental fluorosis), then the EPA derives multiple RfDs (known as candidate RfDs) and then selects a single RfD to protect against all harmful health effects. The RfD can be used to inform public health decisions, including potential actions under the Safe Drinking Water Act.

The objective of dose-response analysis is to derive a toxicity value, called the reference dose (RfD), for total daily oral fluoride intake. An RfD is an estimate of daily oral exposure to the human population (including sensitive populations and lifestages) that is likely to be without appreciable risk of deleterious effects during a lifetime (U.S. EPA, 2002b).

In the future Fluoride Human Health Toxicity Assessment, the dose-response analysis (also referred to as exposure-response analysis)[9] identifies informative biologically and statistically relevant exposure-response data in individual studies that relate changes in fluoride exposure (e.g., fluoride concentrations in drinking water, or biomarkers such as fluoride concentrations in blood or urine) to changes in neurodevelopment or dental fluorosis. The EPA identifies the

[9]Epidemiological studies involve "exposure-response" datasets or effect estimates derived from exposure-response modeling, where exposed populations are studied for a particular health outcome. For example, in the Fluoride Human Health Toxicity Assessment, epidemiological studies relate changes in fluoride exposure measurements (e.g., fluoride concentrations in drinking water or urine) to changes in health outcomes (i.e., developmental neurotoxicity and dental fluorosis). This contrasts with "dose-response" datasets from animal toxicology studies, in which animals are observed for a response after being dosed with a known level of contaminant. The risk assessment paradigm for analyzing this kind of data classically relied on animal toxicology, so the methods are typically referred to as "dose-response analyses." This protocol will generally refer to the datasets as "exposure-response" and the methods as either "exposure-response" or "dose-response."

studies and datasets that are most appropriate for use in dose-response analyses and toxicity value derivation using systematic review methods (U.S. EPA, 2022b) that have been reviewed by the National Academies of Science, Engineering and Medicine (NASEM, 2021).

These studies and datasets are analyzed to determine the exposure-response relationships between fluoride and the harmful health effects that are relevant to the Fluoride Human Health Toxicity Assessment. Within each study or dataset, the EPA identifies a specific fluoride exposure level called the point of departure (POD), which is within or near the observed exposure range and which represents the level below which harmful effects are minimal or unlikely and serves as the beginning point from which to derive the RfD (U.S. EPA, 2012). In this case, this POD is a fluoride concentration (e.g., in drinking water, blood, or urine) and is mathematically converted to an average daily human fluoride intake value (dose in mg/kg/day) (U.S. EPA, 2022b). The POD is then adjusted for potential sources of uncertainty and variability to arrive at an RfD by applying the most appropriate value for each uncertainty factor (i.e., through consideration of the data landscape, potentially sensitive subpopulation(s), etc.).

Candidate RfDs may be derived for multiple health effects (e.g., neurodevelopmental effects, dental fluorosis). A single RfD is selected from among the candidate RfDs to protect against all harmful health effects. The RfD for fluoride derived in the toxicity assessment will represent the daily total fluoride intake a person of any lifestage can consume over a lifetime and be unlikely to experience harmful health effects, including dental fluorosis, neurodevelopmental effects, and other harmful effects.

This section of the protocol provides an overview of the dose-response analysis approach, including considerations for conducting the dose-response analysis, particularly statistical considerations specific to dose-response analysis for toxicity value derivation. These considerations are consistent with EPA guidance. Data requirements and other considerations for dose-response analysis are contained in EPA guidance and support documents, particularly the EPA's *Benchmark Dose Technical Guidance* (U.S. EPA, 2012) and *Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002b) as well as the EPA systematic review methodology (U.S. EPA, 2022b), which was reviewed by the National Academies (NASEM, 2021).

8.1. Selecting Studies and Datasets for Dose-Response Analysis

Section Overview: To begin the dose-response analysis, the EPA uses a well-documented and transparent systematic process to identify the most reliable and informative human studies that contain useable fluoride exposure and response data, based on the study evaluations described in Section 6. The result is a set of studies that represents the best-available science and can be used for toxicity value derivation.

Dose-response analysis begins with a review of the informative studies (i.e., *Low, Medium,* or *High* confidence) in the fluoride literature database to identify those best able to support dose-response analysis and RfD derivation. Preferred studies are those determined to have sufficient

data for dose-response analysis as well as strengths with respect to key attributes (e.g., study confidence, range of exposure, confounding, measurement of the health outcome, study size, modeling approach), as described in the following subsections. Studies meeting the criteria described in this section advance to RfD derivation.

The first consideration for selection of a study for dose-response analysis is whether the study contains exposure-response information that is feasible to analyze. For the Fluoride Human Health Toxicity Assessment, this information might be in the form of exposure-response datasets (e.g., tabulated fluoride measurements and incidence of dental fluorosis) or exposure-response models in publications identified from the systematic review (e.g., effect estimates from regression analysis relating fluoride measurements to changes in IQ). In both instances, the fluoride exposure measures might be external (such as concentrations of fluoride in drinking water) or internal (biomarkers such as fluoride concentration in urine). The requirements for determining whether a given set of exposure-response information is feasible to analyze vary depending on the data types—that is, whether the exposure is treated as categorical (i.e., "low," "medium," or "high;" quantiles) or continuous (i.e., measurements), and similarly, whether the response data are dichotomous (e.g., presence/absence of dental fluorosis) or continuous (e.g., IQ). Table 8-1 provides examples of reporting requirements based on data structure.

Table 8-1. Examples of Data Requirements for Dose-Response Analysis

Exposure		Response	
Categorical	Exposure group[a] median[b], the number of study participants per group, cut points, units, and any transformations	Dichotomous[c]	Number, percent, or proportion of responders per exposure group[d,e]
		Continuous	Outcome mean and standard deviation (or other measures of variance)[d,f]
Continuous	Exposure range (minimum and maximum), units, and any transformations	Dichotomous[c]	Estimate (e.g., risk ratio per *x* units); 95% CI, p-value or equivalent measure; model form (e.g., logistic, log-logistic); transformations; confounders adjusted; and sample size
		Continuous	Slope estimate (per *x* units); 95% CI, p-value or equivalent measure; model form (e.g., linear, log-linear); transformations; confounders adjusted; and sample size

CI = confidence interval.

[a]There should be more than two exposure groups (in addition to the reference or control group) or at least as many exposure groups as dose-response model parameters.

[b]Medians are preferred over means but means may be used. If neither are reported but the bounded range is provided, range midpoints may be considered but are not preferred for dose-response analysis.

[c]The requirements for categorical response data (e.g., multilevel responses like Dean's Index for dental fluorosis) are the same for dichotomous response data.

[d]For multilevel categorical responses (e.g., Dean's Index with mild, moderate, severe), the number, percent or proportion of responders should be reported for each level and exposure group.

[e]Estimates from regressions based on grouped exposure are not suitable for current EPA dose-response analysis methods.

[f]Standard deviations may be back-calculated from confidence intervals and p-values.

Once it is determined that the data meet the minimum requirements for feasibility to perform an analysis, the next step is to assess the degree to which the data support dose-response analysis. Considerations related to this are described in detail in the *Benchmark Dose Technical Guidance* (U.S. EPA, 2012). In general, such data ideally have the following features:

- A biologically meaningful or statistically significant dose-response trend. (Note that statistical significance is not a requirement).
- Responses that are spread (not clustered) across the exposure range.
- Exposure ranges that provide sufficient coverage (i.e., include low, medium, and high levels of exposures), particularly at lower exposure levels. Ideally, the study assesses exposures at or near where any prespecified adverse effect rate thresholds are observed (e.g., the BMR under a benchmark dose modeling framework (Section 8.2.1.2)) to minimize low-dose extrapolations.

Datasets that do not have these features are treated with care, as they are expected to have greater uncertainty (U.S. EPA, 2012). In instances where studies contain their own dose-response analysis (i.e., regression of exposure vs. response) and the EPA is working with those effect estimates, the listed features should be applied to the extracted information describing the dose-response function from the scientific literature.

The combination of datasets via meta-analysis might be considered to overcome individual dataset limitations. Studies with quantitative information might be meta-analyzed if pooling is

reasonable (e.g., similar designs) or if data can be transformed into a common metric (e.g., log scale transformations).

Finally, studies with appropriate exposure-response datasets also undergo an evaluation and comparison based on key study attributes, such as study confidence, exposure paradigm, participant selection, measurement of the health outcome, confounding, and study design (Table 8-2). All study attributes are examined from the perspective of potential impacts on dose-response analysis for a given health outcome. *High*- and *Medium*-confidence studies are considered most useful for dose-response analysis; however, *Low*-confidence studies might be considered, especially if *High*- and *Medium*-confidence studies are not available or have limited dose-response data.

Table 8-2. Examples of Attributes Used to Prioritize Epidemiological Studies for Dose-Response Analysis for the Fluoride Human Health Toxicity Assessment[a]

Study Attributes		Considerations
Study confidence[b]		*High*- or *Medium*-confidence studies are highly preferred over *Low*-confidence studies. The available *High*- and *Medium*-confidence studies are further differentiated based on the study attributes as well as a reconsideration of the specific limitations identified and their potential impacts on dose-response analyses.
Relevance of exposure paradigm	Exposure durations	When developing a chronic toxicity value, chronic or subchronic studies are preferred over studies of acute exposure durations. Exceptions exist, such as when a susceptible population or lifestage is more sensitive in a particular time window (e.g., developmental exposure).
	Exposure levels	Exposures near or within the range of typical environmental human exposures are preferred. Studies with a broad exposure range and multiple exposure levels are preferred to the extent that they can provide information about the shape of the exposure-response relationship (the EPA *Benchmark Dose Technical Guidance*, (U.S. EPA, 2012)) and facilitate extrapolation to more relevant (generally lower) exposures.
Measurement of exposure during relevant developmental window		Studies that can reliably distinguish between levels of exposure within an exposure window considered most relevant for development of a causal effect are preferred. Exposure assessment methods that provide quantitative measurements at the level of the individual and that reduce measurement error are preferred.
Participant selection		Studies that provide risk estimates for the most susceptible groups are preferred.
Control for possible confounding[c]		Studies with a design (e.g., matching procedures, blocking) or analysis (e.g., confounders or other procedures for statistical adjustment) that adequately address the relevant sources of potential critical confounding for a given outcome are preferred.
Measurement of health outcome(s)		Studies that can reliably distinguish the presence or absence (or degree of severity) of the outcome are preferred. Outcome ascertainment methods using generally accepted or standardized approaches are preferred.
		Studies with individual data are preferred. Examples of such studies include those that characterize the overall incidence of individuals affected by relevant outcomes.
		Among several relevant health outcomes, preference is generally given to those with greater biological or potential public health significance.
Study size and design		Preference is given to studies using designs reasonably expected to have power to detect responses of suitable magnitude.[d] This does not mean that studies with substantial responses but low power would be ignored but instead that they should be interpreted in light of a confidence interval or variance for the response. Studies that

Study Attributes	Considerations
	address changes in the number of participants at risk (through decreased survival, loss to follow-up) are preferred.

[a]The table is adapted from Table 7-4 in the EPA's externally peer-reviewed human health systematic review methodology (U.S. EPA, 2022b).
[b]Section 6 provides information about how study confidence is determined during study evaluation.
[c]Confounding refers to a variable that is associated with both exposure and outcome but is not an intermediary between the two.
[d]Power is an attribute of the design and population parameters, based on a concept of repeatedly sampling a population; it cannot be inferred post hoc using data from one experiment (Hoenig and Heisey, 2001).

8.2. Conducting Dose-Response Analyses

Section Overview: In the dose-response analysis, the EPA calculates the amount of fluoride a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (i.e., the RfD). The EPA uses transparent, long-standing, peer-reviewed methods to conduct a rigorous analysis of the set of studies selected as representing the best available science, to identify an exposure level below which harmful effects start to become unlikely (i.e., the point of departure or POD). This level is then adjusted to account for sources of uncertainty and variability (i.e., by applying UFs). The result is a toxicity value derived transparently and systematically that can be used to inform public health decisions and support the Safe Water Drinking Act.

For the future Fluoride Human Health Toxicity Assessment, the EPA performs dose-response analyses using a systematic, two-step process (described below) that is applied to the set of studies and data meeting the criteria for inclusion in dose-response analysis. The result is a toxicity value, called a reference dose (RfD), which represents the amount of fluoride a person can ingest daily over a lifetime and be unlikely to experience harmful health effects (U.S. EPA, 2012, 2005, 2002b).

First, within each study or dataset, a dose-response analysis is carried out in which the EPA analyzes how increasing dose (exposure) impacts harmful health effects (response). From this relationship, the EPA identifies a level of fluoride exposure, called the POD, which is within or near the observed exposure range for a given study or dataset and which represents the level of exposure below which the harmful effects are minimal or not expected and that is used to begin the RfD derivation process (U.S. EPA, 2012). This level is typically a fluoride concentration (e.g., in drinking water, blood, or urine), which is then mathematically converted to an average daily human fluoride intake value (dose) in mg/kg/day (the POD_{ADI}) (U.S. EPA, 2022b).

Second, the POD_{ADI} is adjusted for potential sources of uncertainty and variability. The EPA uses peer-reviewed human health risk assessment methodologies and best available science to make determinations about the level of uncertainty associated with each of five areas of uncertainty and assigns an uncertainty factor (UF) for each. A total or "composite" UF (UF_C) is calculated by multiplying the individual UFs together.

To finally arrive at an RfD, the POD_{ADI} is divided by the UF_C:

$$RfD = POD_{ADI} / UF_C$$

The resulting RfD is an estimate of daily oral fluoride exposure to the human population (including sensitive populations and lifestages) that is likely to be without appreciable risk of deleterious effects during a lifetime, and that accounts for uncertainty to ensure that the EPA is protecting human health based on the best available science.

The EPA might derive RfDs from multiple relevant datasets across both neurodevelopmental and dental fluorosis outcomes. These are considered candidate RfDs and are evaluated collectively before a single overall RfD is selected to protect against all harmful health effects.

The sections that follow describe the dose-response analysis methods used to determine PODs (Section 8.2.1) and how the EPA plans to select and apply UFs (Section 8.2.2).

8.2.1. Estimating Points of Departure

A POD can be one of three types: a no-observed-adverse-effect level (NOAEL), a lowest-observed-adverse-effect level (LOAEL), or the lower bound on a benchmark dose (BMDL) derived from benchmark dose (BMD) modeling. A POD derived from a NOAEL, LOAEL, or BMDL relies on study-specific exposure measurements (e.g., fluoride concentration in drinking water, or a biomarker such as fluoride concentration in urine), expressed in mg/L or ppm. An RfD is defined as daily total oral fluoride exposure and is expressed in mg/kg/day. Therefore, concentration-based PODs are converted into average-daily-intake-based PODs (POD_{ADI}) during RfD derivation (U.S. EPA, 2022b).

The following subsections provide an overview of the NOAEL/LOAEL and BMD modeling approaches, as well as methods for converting to average daily intake, which are consistent with EPA guidance and support documents, especially the *Benchmark Dose Technical Guidance* (U.S. EPA, 2012), *Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002b), and *Approaches for the Application of Physiologically Based Pharmacokinetic (PBPK) Models and Supporting Data in Risk Assessments* (U.S. EPA, 2006).

Any of the frameworks for dose-response analysis (i.e., the NOAEL/LOAEL approach or benchmark dose modeling) might be applied in the future Fluoride Human Health Toxicity Assessment, depending on the available datasets. All dose-response analysis and PK modeling choices, assumptions, and sensitivity analyses will be documented in the toxicity assessment to ensure transparency.

8.2.1.1. NOAEL or LOAEL Approach

The NOAEL or LOAEL method is a non-modeling approach for determining the POD. The NOAEL is the maximum observed non-zero exposure level at which no biologically meaningful increase in occurrence (or severity) of an adverse effect is observed in a study. Exposures at or below a NOAEL are considered likely to be without appreciable risk, and therefore NOAELs might be used as PODs.

If a NOAEL is not observed (i.e., there is no non-zero exposure level without an observed adverse effect captured by the study design), then the POD is set to the lowest exposure level, the LOAEL. Since the LOAEL is not representative of an exposure level at which no adverse effects

are observed, a UF is automatically applied when a LOAEL is used as the POD (Section 8.2.2 for details).

A limitation of this approach is that it does not use the full information in the dataset (e.g., variability, shape of the exposure-response curve). Other limitations exist and are well-documented (U.S. EPA, 1995; Leisenring and Ryan, 1992; Kimmel and Gaylor, 1988; Crump, 1984; Gaylor, 1983). Because of these limitations, PODs are only derived using the NOAEL/LOAEL approach when the available data are not suitable for benchmark dose modeling (described in the following section).

8.2.1.2. Benchmark Dose Modeling

BMD modeling empirically characterizes the dose-response relationship from observed data and estimates the exposure level (the BMD) that is associated with a prespecified amount of change in the outcome (known as the benchmark response or BMR). A lower confidence limit is placed on the BMD to obtain a dose (BMDL) that assures with a high confidence level (e.g., 95%) that the BMR is not exceeded (U.S. EPA, 2012). Use of the BMDL as the POD accounts for uncertainty and variability due to study design (U.S. EPA, 2012).

This following subsection describes considerations in choosing the BMR, dose-response functions and modeling, frequentist and Bayesian frameworks for inference, and other BMD modeling approaches.

8.2.1.2.1. The Benchmark Response

The BMR represents the magnitude of the change in a health effect that is considered adverse. The BMR is selected based on agency guidance and the biological and statistical characteristics of the dataset as well as the applications for which the resulting POD will be used, as different applications might warrant different BMR values (U.S. EPA, 2012). BMR selection involves weighing biological and statistical considerations such as the severity of the endpoint and the observable range of response.

Various types of BMRs exist, depending on the type of outcome being analyzed (i.e., dichotomous or continuous; or continuous data that is dichotomized).

For dichotomous outcomes, the BMR could be defined in terms of added or extra risk. Added risk is defined as the increase in the chance of an adverse outcome associated with an exposure. Extra risk is the added risk after accounting for cases that would have occurred in the absence of the exposure (the background risk) and is preferred for its flexibility in accounting for non-zero background rates.[10] A BMR of 10% extra risk is classically recommended as a standard reporting level and for minimally adverse effects because it is near the low end of the observable range for many common experimental animal study designs (U.S. EPA, 2012). A BMR of 5% extra risk or lower is typically used for more severe effects (e.g., developmental effects) or for some epidemiological data. Extra risk BMRs of 1% or lower are most often used for epidemiological data because these studies are typically based on larger sample sizes that have

[10]Extra risk simplifies to added risk if the data support a background rate of zero. By choosing to work with extra risk, assumptions about the presence or absence of background levels might be avoided and instead, informed by the data.

more power to estimate smaller extra risks, and/or due to severity of the outcome (e.g., IQ decrements) (U.S. EPA, 2025a, 2024g, 2023b, d, 2022a).

For continuous outcomes, the BMR is ideally based on an established definition of the level of change, which is biologically significant. In the absence of any such definition, the BMR is often defined as one standard deviation from the control mean for minimally adverse effects, and for more severe effects, one half-standard deviation is often used.[11] The BMR could also be defined as absolute or relative deviation from the mean.

Continuous outcomes (e.g., IQ decrements) can also be analyzed using dichotomous BMRs, e.g., extra risk, if they are dichotomized, or converted, into binary data based on a known threshold of adversity. For example, continuous measures of birth weight could be converted into dichotomous realizations of low birth weight (below 2500 grams) and analyzed as dichotomous data. Because of the "combination" of continuous and dichotomous data, this approach is referred to as "hybrid" benchmark dose analysis (Section 8.2.1.2.4 and Appendix E).

For both continuous and dichotomous datasets, the BMR might be mathematically expressed as a function of the difference between the modeled responses at the benchmark dose and at background levels. For example, if the response is severe dental fluorosis, represented as a dichotomous data type (i.e., measured as the presence or absence of severe dental fluorosis), then the BMR might be based on extra risk, which has the following form:

$$BMR = \frac{p(d_{BMD}) - p(d_0)}{1 - p(d_0)}$$

where $p(d)$ is a mathematical equation describing how dental fluorosis incidence changes with exposure to fluoride, or dose, d. Therefore, $p(d_{BMD})$ represents the incidence at the benchmark dose, and $p(d_0)$ represents the background incidence when $d = 0$. Conceptually, the BMD is found by specifying the exposure-response function $p(d)$, using inference to estimate the model parameters, then solving the equation for the BMD and subsequently calculating the BMDL.

8.2.1.2.2. Dose-Response Functions

A variety of functional forms define the dose-response relationship, $p(d)$, and the EPA has developed a standard set of models, including linear and nonlinear options, which can be applied to datasets to characterize dose-response relationships. Alternative models could be explored in instances where biological support is substantial; all alternatives should be presented with clearly outlined strengths as well as any limitations. All potentially appropriate models are considered candidates and are evaluated and compared based on success metrics like goodness-of-fit statistics and other dose-response analysis-specific criteria. Inference is then often based on the single model exhibiting the strongest support under the chosen criterion.[12]

[11]For normalized continuous data (i.e., continuous data transformed to have the distribution of a standard normal data with a mean of zero and a standard deviation of one), the BMR of one control standard deviation from the control mean is comparable to a BMR of 10% extra risk, which is considered a standard reporting level (U.S. EPA, 2012).

[12]Alternatives exist to single-model selection, such as Bayesian model averaging. This technique produces an estimate from multiple models simultaneously by leveraging contributions from every candidate model and avoiding the need to choose a single, "best" model. This is discussed in detail later in Section 8.2.1.2.3.

The EPA has developed guidance for assessing model fit, selecting suitable models, and reporting model comparison results in the *Benchmark Dose Technical Guidance* (U.S. EPA, 2012) and the EPA's *Benchmark Dose Software (BMDS) Online User Guide*. Alternative analyses are used if the procedure fails to yield reliable results; for example, if the fit is poor, modeling could be restricted to the lower doses, especially if there is competing toxicity at higher doses.

8.2.1.2.3. Frameworks for Inference: Frequentist and Bayesian

Inference from dose-response models about the BMDL could be drawn using either frequentist or Bayesian frameworks (Table 8-3). The former involves maximum likelihood estimation and relies only on the observed data to produce point estimates for the BMD (representing the most likely value given the observed data) and the BMDL. Models are assessed for goodness-of-fit using various metrics, and model comparison is carried out on the basis of the Akaike Information Criterion (AIC). This frequentist approach has been used routinely at the EPA and is documented in U.S. EPA (2012).

Bayesian frameworks allow the addition of outside knowledge to the observed exposure-response data. In 2014, the NRC recommended that the EPA adopt Bayesian dose-response frameworks, citing advantages of POD estimates that incorporate important knowledge outside of the critical dose-response study (or studies) (e.g., mechanistic study findings, dose-response patterns for similar chemicals, known measurement error in exposure levels) (NRC, 2014). Additionally, probabilistic estimates of the BMD could be integrated with PK models that consider population-specific physiological variations, enabling the calculation of average daily intake while addressing inherent human-specific PK uncertainty.

In a Bayesian framework, any relevant outside information is quantitatively represented with a "prior" probability distribution (i.e., a range of plausible values), which is then formally combined with observations to generate a "posterior" probability distribution for the BMD. This posterior distribution represents a range of plausible values for the BMD, considering both the observed data and relevant outside information. The median of the posterior distribution is taken to be the BMD estimate, and the BMDL is set to the lower bound of the BMD's credible interval.[13]

Under the Bayesian framework, model fit and comparison could be carried out using various metrics and criteria (e.g., Bayesian Information Criteria (BIC)) to select a single best model. Alternatively, Bayesian model averaging could be used in lieu of selecting a single model, to leverage the strengths of each model in estimating the POD. In Bayesian model averaging, priors are specified for each candidate model, the data are used to obtain posterior model probabilities, and quantities of interest (e.g., the BMD) are combined across models with weights given by those probabilities. In this way, models with greater posterior support contribute more while

[13]A credible interval is the Bayesian equivalent of a confidence interval, in that it may be interpreted as a range of values in which the analyst may have some confidence that the true BMD is included. However, a credible interval has the added benefit of being able to represent a valid probability. In other words, a BMDL from a credible interval represents the number above which the true BMD lies with a certain level of probability (for example, 90% or 95% probability).

other candidate models might still contribute, thereby explicitly accounting for model uncertainty and reducing sensitivity to any one model specification (Wheeler et al., 2019; Wheeler and Bailer, 2009).

The EPA's Benchmark Dose Software currently provides modeling options and documentation on dichotomous Bayesian modeling and Bayesian model averaging (*BMDS Online User Guide*). Further, the EPA has published assessments that leverage Bayesian methods including Bayesian PK meta-analysis (U.S. EPA, 2025d, 2024h, 2023d) and Bayesian dose-response meta-analysis (U.S. EPA, 2025c). Bayesian analysis and model averaging might be applied in the dose-response analysis for the future Fluoride Human Health Toxicity Assessment, depending on the available data sets.

Table 8-3. Benchmark Dose Modeling Under Frequentist and Bayesian Frameworks

	Frequentist	Bayesian
BMD	Single estimate (maximum likelihood estimate)	Full probability distribution (the "posterior"); summarize by the median
BMDL	Lower confidence bound on the BMD	Lower credible bound on the BMD
Evidence Used	Observed study data	Observed study data plus relevant external information (e.g., mechanistic data, measurement error)
Model Fit and Comparison	Select a single "best" model using model fit criteria (e.g., AIC)	Select a single "best" model using model fit criteria (e.g., Bayesian Information Criteria: BIC), plus the option to employ model averaging and reduce sensitivity to any single model
EPA Guidance	The EPA's *Benchmark Dose Technical Guidance (2012)*	The *BMDS Online User Guide* covers Bayesian dichotomous BMD modeling and model averaging. Bayesian methods were also used in recent EPA risk assessments (U.S. EPA, 2025c)

BMD = benchmark dose; BMDL = benchmark dose lower confidence limit; AIC = Akaike Information Criteria.

8.2.1.2.4. Hybrid Benchmark Dose Analysis

In some cases, biological and statistical evidence indicates that a continuous response level above (or below) a certain threshold, or cutoff value, may be considered adverse. In these situations, a continuous outcome is dichotomized (i.e., transformed) to represent being above the known threshold of adversity. To do this, each observed measurement is re-considered as a "1" if it is above the threshold, and a "0" if it is below, resulting in a new, dichotomous dataset. Accordingly, the BMR is chosen for categorical data and defined under a dichotomous dose-response analysis framework (Section 8.2.1.2.1). Because of the "combination" of continuous and dichotomous data types, this approach is referred to as hybrid benchmark dose analysis.

For example, if an exposure causes decreased birth weight (a continuous outcome), the modeler would ascertain whether there is an established definition of how much of a decrease in birth weight is considered adverse (i.e., a threshold of adversity). For epidemiological studies, low birthweight is defined as less than 2,500 grams. In this case, the modeler would use the hybrid method, dichotomize the data based on a threshold of 2,500 grams, and set the BMR at 5% extra risk to represent a moderate adverse outcome of being a low birthweight baby in an exposed population.

In summary, when working with continuous outcomes, the modeler has the choice of carrying out a continuous dose-response analysis (i.e., choosing continuous BMRs, continuous dose-response models), or dichotomizing the data so that it represents being above or below a known threshold of adversity (if such a threshold exists). Specifically, for a BMD analysis of continuous data like IQ decrements, either the absolute deviation BMR or hybrid approach with extra risk BMR could be used to derive a POD.

The hybrid approach can also be used as a framework to analyze observational results from a study that already includes dose-response analysis. For example, if the study reports an effect estimate and standard errors from a regression analysis, the hybrid approach defines a BMR for continuous endpoints by assigning an adversity cutoff value to the underlying distribution, allowing the BMD to be estimated as an added or extra risk of an individual falling into the adverse tail of the population. Appendix E provides more details.

8.2.1.3. Dose-Response Meta-Analysis

PODs might be study-specific; alternatively, when several studies of the same outcome can be reasonably combined or transformed to a common scale, a pooled POD from a meta-analysis could also be considered as a candidate. For example, the EPA's *Toxicological Review of Inorganic Arsenic* used hierarchical Bayesian dose-response meta-analysis to synthesize data from multiple epidemiologic studies to derive a health-based candidate RfD (U.S. EPA, 2025c).

8.2.1.4. Deriving Average Daily Fluoride Intake PODs

The estimated POD (e.g., BMDL, NOAEL, or LOAEL) is based on measured fluoride exposure (in drinking water or biological samples like blood or urine) linked to an observed health outcome in a study population. As described in the Section 8 introduction, the RfD is expressed as an average daily human fluoride intake (in mg/kg/day), and any exposure-based POD must first be converted to a corresponding fluoride average daily intake POD (i.e., the POD_{ADI}). This conversion to intake (i.e., dose) accounts for fluoride exposure from drinking water and non-drinking water sources, as well as variability between populations in water ingestion and the proportion of exposure from drinking water. For chronic effects, average daily intake is averaged over the life span while average daily intake during a critical period is averaged over the appropriate critical window of exposure (U.S. EPA, 2005, 1991)..

To make this conversion, a PK analysis is used. As outlined in Section 6.3, PK analyses, including PBPK modeling, quantify the relationship between total applied dose and measurable biomarkers, such as concentrations in the blood or urine. Once this mathematical relationship is established, these models can be run in the "reverse" direction to infer the applied dose from a biomarker concentration. The half-life of fluoride in human serum has been estimated to be between 3 and 10 hours (NTP, 2016) with most fluoride excreted in the urine and minimal excretion through feces (Waterhouse et al., 1980). Hence, urinary fluoride characterizes both recent and long-term exposure, as it reflects intake from the past one or two days as well as accumulated fluoride in skeletal tissue from chronic exposure that is released back into circulation. The extent to which urinary fluoride concentration reflects either recent or long-term exposure depends on factors such as the timing of urination following exposure, along with lifestage-specific factors affecting both urinary excretion and bone uptake. Therefore, PK

analyses can be used to inform the fraction of ingested fluoride expected to be present in the urine and use lifestage-specific parameters to estimate the fluoride intake based on a given urinary concentration. An even more robust analysis relying on a PBPK model can be used to integrate an individual's previous exposure and estimate uptake by and resorption from bone, resulting in a urinary concentration prediction that is more reflective of a specific exposure scenario. Use of a PBPK model in this way would require assumptions about an individual's previous exposure, which can be informed by study-specific information on details such as exposure history, the individual's body weight, total urinary exposure, and time since last ingestion and urination. With these inputs, a PBPK model can be used to estimate the average daily fluoride intake that is consistent with a biomarker POD. In the absence of study-specific data on, for example, intake rates or body weight, the EPA has developed recommended values for use in dose-response analysis (U.S. EPA, 1988).

When a POD is established based on fluoride drinking water concentrations reflecting modern-day exposure patterns (i.e., which includes population exposure to fluoride from non-drinking water sources), it is necessary to obtain study-specific information regarding additional source contributions to estimate total daily intake from all fluoride sources. By incorporating data from these sources, the resulting POD_{ADI} accurately reflects the POD at which total fluoride exposure leads to the observed effect.

8.2.2. Uncertainty Factors and Deriving the Reference Dose

In the dose-response analysis, EPA first identifies the POD, i.e., the exposure level below which harmful effects are minimal or unlikely, that serves as the starting point for the derivation of the RfD.[14] The next step is to apply UFs to account for uncertainty[15] and variability.[16] UFs are numeric values, typically 1, 3, or 10 (i.e., 10^0, $10^{1/2}$ or 10^1), that represent increasing magnitudes of uncertainty arising from specific assumptions or modeling decisions made while deriving PODs. The application of UFs is informed by the agency's peer-reviewed risk assessment methodology (U.S. EPA, 2002b).

Five standard UFs account for separate types of uncertainty commonly encountered in risk assessment, and are considered on an individual, independent basis:

[14]A point of departure is also sometimes defined as "marking the beginning of the extrapolation to the RfD." This is because while a POD falls within or near the range of observed exposure, the application of uncertainty factors adjusts the POD to a lower, presumably more protective level, possibly below the observed range of exposure.
[15]Uncertainty occurs because of a lack of knowledge (U.S. EPA, 2002b). It is not the same as variability. For example, a risk assessor may be very certain that different people drink different amounts of water but may be uncertain about how much variability there is in water intakes within the population. Uncertainty can often be reduced by collecting more and better data, while variability is an inherent property of the population being evaluated. Variability can be better characterized with more data, but it cannot be reduced or eliminated. Efforts to clearly distinguish between variability and uncertainty are important for both risk assessment and risk characterization (U.S. EPA, 2002b).
[16]Variability refers to true heterogeneity or diversity (U.S. EPA, 2002b). For example, among a population that drinks water from the same source and with the same contaminant concentration, the risks from consuming the water may vary. This may be due to differences in exposure (i.e., different people drinking different amounts of water and having different body weights, different exposure frequencies, and different exposure durations) as well as differences in response (e.g., genetic differences in resistance to a chemical dose). Those inherent differences are referred to as variability. Differences among individuals in a population are referred to as inter-individual variability, while differences for one individual over time is referred to as intra-individual variability (U.S. EPA, 2002b).

1. An interspecies uncertainty factor, UF_A, to account for animal-to-human extrapolation when the POD is based on animal data, consisting of equal parts representing toxicokinetic and toxicodynamic differences. If results from animal experiments are not used to make inferences about humans, the UF_A is assigned a value of 1.
2. An intraspecies uncertainty factor, UF_H, to account for variations in susceptibility across the human population, and the possibility that the POD might not be representative of individuals who are most susceptible. The UF_H could be lowered, or set to 1, if data-based adjustments are feasible (U.S. EPA, 2014b). For example, data-based adjustments might include Bayesian modeling to characterize population variability, and toxicodynamic or pharmacokinetic modeling to characterize internal dose (NRC, 2014; U.S. EPA, 2014b, 2002b).[17] [18] The UF_H could also be reduced if the POD is derived from or adjusted specifically for susceptible individuals (i.e., not for a general population that includes both susceptible and non-susceptible individuals) ((U.S. EPA, 2002b), §4.4.5; (U.S. EPA, 1998), §4.2; (U.S. EPA, 1996a), §4; (U.S. EPA, 1994), §4.3.9.1; (U.S. EPA, 1991), §3.4). When data-based adjustments or modeling is not supported, a UF_H with a value of 10 is typically considered.
3. A LOAEL uncertainty factor, UF_L, when the POD is based on a LOAEL instead of a NOAEL or BMDL, to adjust the POD to an exposure level where effects are not expected. If a POD is based on NOAEL or BMDL, a UF_L of 1 may be applied. Otherwise, a UF_L of 3 or 10 could be used depending on the level of available evidence at lower exposures, as well as the magnitude and nature of the response and the shape of the dose-response curve (U.S. EPA, 2002b, 1998, 1996a, 1994, 1991).
4. A subchronic-to-chronic uncertainty factor, UF_S, to account for the uncertainty in using subchronic exposure studies to make inferences about lifetime exposure, and to consider whether lifetime exposure would result in effects at lower exposure levels (e.g., for studies other than subchronic studies). If the exposure is within an etiologically relevant window, modeled or observed, a UF_S of 1 may be applied. Otherwise, a UF_S of 3 or 10 could be applied to the POD, depending on the exposure duration of the studies and the nature of the response (U.S. EPA, 2002b, 1998, 1994).
5. A database uncertainty factor, UF_D, to account for database deficiencies in the available literature for the health effects being assessed (i.e., neurodevelopmental toxicity and dental fluorosis) (U.S. EPA, 2002b, 1998, 1996a, 1994, 1991). The size of the factor depends on the nature of the database deficiency. For example, when deriving RfDs for developmental outcomes, the EPA typically follows the recommendation that a factor of 10 be applied if both a prenatal toxicity study and a two-generation reproduction study

[17]Examples of adjusting the pharmacokinetic portion of interhuman variability include the EPA Toxicological Review of boron and the use of nonchemical-specific kinetic data (e.g., glomerular filtration rate in pregnant humans as a surrogate for boron clearance (U.S. EPA, 2004)) and the EPA Toxicological Review of trichloroethylene and its use of population variability in trichloroethylene metabolism, via a PBPK model, to estimate the lower 1st percentile of the dose metric distribution for each POD (U.S. EPA, 2011c).

[18]Note that when a PBPK model is available for relating human internal dose to environmental exposure, relevant portions of this UF might be more usefully applied prior to animal-to-human extrapolation, depending on the correspondence of any nonlinearities (e.g., saturation levels) between species.

are missing and a factor of $10^{1/2}$ (i.e., 3) if either one or the other is missing (U.S. EPA, 2002b).

Once individual UFs have been identified, the adjustment to the POD takes place in two steps. First, a composite uncertainty factor, UF_C, is calculated by multiplying all individual uncertainty factors together:

$$UF_C = UF_A \times UF_H \times UF_L \times UF_S \times UF_D$$

Note that if any individual UF is assigned a value of 1, it does not contribute to the UF_C, nor the RfD. Also, U.S. EPA (2002b) notes that a UF_C that exceeds 3,000 represents excessive uncertainty and recommends against relying on the resulting RfD.

Next, the POD is divided by the UF_C, reducing the value of the POD in proportion to the magnitude of uncertainty in the analysis. For fluoride, the RfD will be based on the POD derived for average daily intake, described earlier, i.e., the POD_{ADI}. The result is the RfD:

$$RfD = POD_{ADI} / UF_C$$

9. References

Acharya, S. (2005). Dental caries, its surface susceptibility and dental fluorosis in South India. Int Dent J 55: 359-364. https://hero.epa.gov/reference/13029045/

American Dental Association. (2025). Letter to the EPA administrator. https://hero.epa.gov/reference/13283240/

American Public Health Association. (1946). Fluoride. In Standard methods for the examination of water and sewage (9th ed.). New York: American Public Health Association. https://hero.epa.gov/reference/13257276/

ATSDR (Agency for Toxic Substances and Disease Registry). (2003). Toxicological profile for fluorides, hydrogen fluoride, and fluorine [ATSDR Tox Profile]. (NTIS PB2004-100002). Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service. https://hero.epa.gov/reference/192114/

Axelrad, DA; Bellinger, DC; Ryan, LM; Woodruff, TJ. (2007). Dose-response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data. Environ Health Perspect 115: 609-615. https://hero.epa.gov/reference/682395/

Barberio, AM; Hosein, FS; Quiñonez, 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://hero.epa.gov/reference/10488106/

Barberio, AM; Quiñonez, 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: e229-e239. https://hero.epa.gov/reference/8403511/

Bartels, M; Rietveld, MJH; Van Baal, GCM; Boomsma, DI. (2002). Genetic and environmental influences on the development of intelligence. Behavior Genetics 32: 237-249. https://hero.epa.gov/reference/13283207/

Bartos, M; Gumilar, F; Gallegos, CE; Bras, C; Dominguez, S; Cancela, LM; Minetti, A. (2019). Effects of perinatal fluoride exposure on short- and long-term memory, brain antioxidant status, and glutamate metabolism of young rat pups. Int J Toxicol 38: 405-414. https://hero.epa.gov/reference/6796771/

Bartos, M; Gumilar, F; Gallegos, CE; Bras, C; Dominguez, S; Mónaco, 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 α7 nicotinic receptor and oxidative stress. Reprod Toxicol 81: 108-114. https://hero.epa.gov/reference/4985205/

Bashash, M; Thomas, D; Hu, H; Martinez-Mier, EA; Sánchez, BN; Basu, N; Peterson, KE; Ettinger, AS; Wright, R; Zhang, Z; Liu, Y; Schnaas, L; Mercado-García, A; Téllez-Rojo, MM; Hernández-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: 097017. https://hero.epa.gov/reference/4856649/

Bassin, EB; Wypij, D; Davis, RB; Mittleman, MA. (2006). Age-specific fluoride exposure in drinking water and osteosarcoma (United States). Cancer Causes and Control 17: 421-428. https://hero.epa.gov/reference/6408711/

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. https://hero.epa.gov/reference/4932565/

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.

https://hero.epa.gov/reference/452940/
Bradley, RH; Caldwell, BM. (1980). The relation of home environment, cognitive competence, and IQ among males and females. Child Development 1140. https://hero.epa.gov/reference/13283251/
Brochot, C; Toth, J; Bois, FY. (2005). Lumping in pharmacokinetics. J Pharmacokinet Pharmacodyn 32: 719-736. https://hero.epa.gov/reference/818804/
Campbell, JL, Jr.; Clewell, RA; Gentry, PR; Andersen, ME; Clewell, HJ, III. (2012). Physiologically based pharmacokinetic/toxicokinetic modeling. Methods Mol Biol 929: 439-499. https://hero.epa.gov/reference/1478197/
Castiblanco-Rubio, GA; Martinez-Mier, EA. (2022). Fluoride metabolism in pregnant women: a narrative review of the literature [Review]. Metabolites 12. https://hero.epa.gov/reference/13033614/
CDC (Centers for Disease Control and Prevention). (2019). Data quality evaluation of the dental fluorosis clinical assessment data from the national health and nutrition examination survey. (DHHS Publication No. 2019–1383). Hyattsville, Maryland: Centers for Disease Control and Prevention. https://hero.epa.gov/reference/13257199/
CDC (Centers for Disease Control and Prevention). (2024). About community water fluoridation. https://hero.epa.gov/reference/13283241/
Charugundla, BR; Anjum, S; Mocherla, M. (2015). Comparative effect of fluoride, essential oil and chlorhexidine mouth rinses on dental plaque and gingivitis in patients with and without dental caries: a randomized controlled trial. Int J Dent Hyg 13: 104-109. https://hero.epa.gov/reference/10481586/
Chi, DoL; Guinn, Sh; Shands, MaE; Nemawarkar, De; Hill, CoM; Mayhle, Mi; Do, ThT; Li, So; Panchal, Sh. (2024). Socioeconomic vulnerability and access to community water fluoridation in Washington. Journal of the American Dental Association (1939) 155: 747-754.e742. https://hero.epa.gov/reference/13283208/
Chiu, WA; Barton, HA; Dewoskin, RS; Schlosser, P; Thompson, CM; Sonawane, B; Lipscomb, JC; Krishnan, K. (2007). Evaluation of physiologically based pharmacokinetic models for use in risk assessment [Review]. Journal of Applied Toxicology 27: 218-237. https://hero.epa.gov/reference/596339/
Clark, LH; Setzer, RW; Barton, HA. (2004). Framework for evaluation of physiologically-based pharmacokinetic models for use in safety or risk assessment. Risk Analysis 24: 1697-1717. https://hero.epa.gov/reference/818951/
Clarkson, J. (1989). Review of terminology, classifications, and indices of developmental defects of enamel. Adv Dent Res 3: 104-109. https://hero.epa.gov/reference/13088079/
Corrêa, ML; Soares, PSM; da Silva, BGC; Wehrmeister, F; Horta, BL; Menezes, AMB. (2021). Maternal smoking during pregnancy and intelligence quotient in offspring: a systematic review and meta-analysis. NeuroToxicology 85: 99-114. https://hero.epa.gov/reference/12064443/
Crump, KS. (1984). A new method for determining allowable daily intakes. Toxicological Sciences 4: 854-871. https://hero.epa.gov/reference/66651/
Das, TK; Susheela, AK; Gupta, IP; Dasarathy, S; Tandon, RK. (1994). Toxic effects of chronic fluoride ingestion on the upper gastrointestinal tract. J Clin Gastroenterol 18: 194-199. https://hero.epa.gov/reference/12170827/
de Ruiter, R; Botman, E; Teunissen, B; Lammertsma, A; Boellaard, R; Raijmakers, P; Schwarte, L; Nieuwenhuijzen, J; Gonzalez Trotter, D; Eekhoff, E; Yaqub, M. (2024). Performance of simplified methods for quantification of [18F]NaF uptake in fibrodysplasia ossificans progressiva. Frontiers in Nuclear Medicine 4. https://hero.epa.gov/reference/13034318/
Dean, HT. (1942). The investigation of physiological effects of the epidemiological method. American Association for the Advancement of Science 23-31. https://hero.epa.gov/reference/10521195/

Dean, HT; Arnold, FA, Jr; Elvove, E. (1942). Domestic water and dental caries: V. Additional studies of the relation of fluoride domestic waters to dental caries experience in 4,425 white children, aged 12 to 14 years, of 13 cities in 4 states. Public Health Rep 57: 1155-1179. https://hero.epa.gov/reference/10519160/

DenBesten, PK. (1999). Biological mechanisms of dental fluorosis relevant to the use of fluoride supplements. Community Dent Oral Epidemiol 27: 41-47. https://hero.epa.gov/reference/13078164/

DenBesten, PK; Thariani, H. (1992). Biological mechanisms of fluorosis and level and timing of systemic exposure to fluoride with respect to fluorosis. J Dent Res 71: 1238-1243. https://hero.epa.gov/reference/13068896/

Dombrowski, PM; Kakarla, P; Caldicott, W; Chin, Y; Sadeghi, V; Bogdan, D; Barajas-Rodriguez, F; Chiang, SYD. (2018). Technology review and evaluation of different chemical oxidation conditions on treatability of PFAS. Remediat 28: 135-150. https://hero.epa.gov/reference/10178178/

Duckworth, RM; Morgan, SN. (1991). Oral fluoride retention after use of fluoride dentifrices. Caries Res 25: 123-129. https://hero.epa.gov/reference/3999159/

Edmunds, WM; Smedley, PL. (2012). Fluoride in natural waters. In Essentials of medical geology: Revised edition. Netherlands: Dordrecht: Springer. https://hero.epa.gov/reference/13283242/

EFSA Scientific Committee. (2020). Draft framework for protocol development for EFSA's scientific assessments. EFSA Supporting Publications 17: 1843E. https://hero.epa.gov/reference/13283256/

EFSA Scientific Committee. (2025). Updated consumer risk assessment of fluoride in food and drinking water including the contribution from other sources of oral exposure. EFSA journal, 23: e9478. https://hero.epa.gov/reference/13037056/

Ekstrand, J; Alván, G; Boréus, LO; Norlin, A. (1977). Pharmacokinetics of fluoride in man after single and multiple oral doses. Eur J Clin Pharmacol 12: 311-317. https://hero.epa.gov/reference/13033637/

Ekstrand, J; Boreus, LO; de Chateau, P. (1981). No evidence of transfer of fluoride from plasma to breast milk. Br Med J 283: 761-762. https://hero.epa.gov/reference/19721/

Ekstrand, J; Ehrnebo, M. (1979). Influence of milk products on fluoride bioavailability in man. Eur J Clin Pharmacol 16: 211-215. https://hero.epa.gov/reference/13033944/

Ekstrand, J; Ehrnebo, M. (1983). The relationship between plasma fluoride, urinary excretion rate and urine fluoride concentration in man. J Occup Med 25: 745-748. https://hero.epa.gov/reference/13033425/

Ekstrand, J; Spak, CJ. (1990). Fluoride pharmacokinetics: its implications in the fluoride treatment of osteoporosis. J Bone Miner Res 5 Suppl 1: S53-S61. https://hero.epa.gov/reference/13032161/

Elekdag-Turk, S; Almuzian, M; Turk, T; Buzalaf, MAR; Alnuaimi, A; Dalci, O; Darendeliler, MA. (2019). Big toenail and hair samples as biomarkers for fluoride exposure - a pilot study. BMC Oral Health 19: 82. https://hero.epa.gov/reference/10481992/

Elvove, E; Goudey, RF. (1941). Comments on committee report on methods of determining fluorides. Journal of the American Water Works Association 2210-2214. https://hero.epa.gov/reference/13257286/

Erickson, JD. (1980). Down syndrome, water fluoridation, and maternal age. Teratology 21: 177-180. https://hero.epa.gov/reference/12170744/

Erickson, JD; Oakley, GP, Jr.; Flynt, JW, Jr.; Hay, S. (1976). Water fluoridation and congenital malformations: no association. J Am Dent Assoc 93: 981-984. https://hero.epa.gov/reference/12170668/

Erickson, KI; Hillman, C; Stillman, CM; Ballard, RM; Bloodgood, B; Conroy, DE; Macko, R; Marquez,

DX; Petruzzello, SJ; Powell, KE. (2019). Physical activity, cognition, and brain outcomes: a review of the 2018 physical activity guidelines. Medicine & Science in Sports & Exercise 51. https://hero.epa.gov/reference/13283211/

Ershow, AB; Cantor, KP. (1989). Total water and tapwater intake in the united states: population-based estimates of quantities and sources. Bethesda, MD: Life Sciences Research Office. https://hero.epa.gov/reference/710071/

Eskandari, F; Kumah, EA; Azevedo, L; Stephenson, J; John, S; Zohoori, FV. (2023). Fluoride exposure in community prevention programmes for oral health using nail clippings and spot urine samples: a systematic review and meta-analysis. Caries Res 57: 197-210. https://hero.epa.gov/reference/13059195/

Frost, M; Blake, G; Cook, GJ; Marsden, P; Fogelman, I. (2009). Differences in regional bone perfusion and turnover between lumbar spine and distal humerus: (18)F-fluoride PET study of treatment-naive and treated postmenopausal women. Bone 45. https://hero.epa.gov/reference/13034262/

Gaylor, DW. (1983). The use of safety factors for controlling risk. J Toxicol Environ Health 11: 329-336. https://hero.epa.gov/reference/63924/

González-Horta, C; Ballinas-Casarrubias, L; Sánchez-Ramírez, B; Ishida, MC; Barrera-Hernández, A; Gutiérrez-Torres, D; Zacarias, OL; Saunders, RJ; Drobná, Z; Mendez, MA; García-Vargas, G; Loomis, D; Stýblo, M; Del Razo, LM. (2015). A concurrent exposure to arsenic and fluoride from drinking water in Chihuahua, Mexico. Int J Environ Res Public Health 12: 4587-4601. https://hero.epa.gov/reference/2854618/

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 173: 940-948. https://hero.epa.gov/reference/5930232/

Green, R; Rubenstein, J; Popoli, R; Capulong, R; Till, C. (2020a). Sex-specific neurotoxic effects of early-life exposure to fluoride: a review of the epidemiologic and animal literature. Curr Epidemiol Rep 7: 263-273. https://hero.epa.gov/reference/10486120/

Green, R; Till, C; Cantoral, A; Lanphear, B; Martinez-Mier, EA; Ayotte, P; Wright, RO; Tellez-Rojo, MM; Malin, AJ. (2020b). Associations between urinary, dietary, and water fluoride concentrations among children in Mexico and Canada. Toxics 8. https://hero.epa.gov/reference/10490281/

Hawkins, RA; Choi, Y; Huang, SC; Hoh, CK; Dahlbom, M; Schiepers, C; Satyamurthy, N; Barrio, JR; Phelps, ME. (1992). Evaluation of the skeletal kinetics of fluorine-18-fluoride ion with PET. J Nucl Med 33: 633-642. https://hero.epa.gov/reference/13034146/

Health Canada. (2010). Guidelines for Canadian drinking water quality: guideline technical document – fluoride. (ISBN 978-1-100-18261-2). Ottawa, Ontario: Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada. https://hero.epa.gov/reference/10528541/

Health Canada. (2024). Expert panel meeting on the health effects of fluoride in drinking water: summary report, June 8-9, 2023. Ottawa, ON: Health Canada. https://hero.epa.gov/reference/13246052/

Hefferon, R; Goin, DE; Sarnat, JA; Nigra, AE. (2024). Regional and racial/ethnic inequalities in public drinking water fluoride concentrations across the US. J Expo Sci Environ Epidemiol 34: 68-76. https://hero.epa.gov/reference/11357714/

Heindel, JJ; Bates, HK; Price, CJ; Marr, MC; Myers, CB; Schwetz, BA. (1996). Developmental toxicity evaluation of sodium fluoride administered to rats and rabbits in drinking water. Fundam Appl Toxicol 30: 162-177. https://hero.epa.gov/reference/4389758/

Higgins, JPT; Green, S. (2011). Cochrane handbook for systematic reviews of interventions. version 5.1.0 (updated March 2011). London, UK: The Cochrane Collaboration.

https://hero.epa.gov/reference/3507864/
Hodge, HC. (1950). The concentration of fluorides in drinking water to give the point of minimum caries with maximum safety. J Am Dent Assoc 40: 436-439. https://hero.epa.gov/reference/2528292/
Hoenig, JM; Heisey, DM. (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. Am Stat 55: 19-24. https://hero.epa.gov/reference/713704/
Hong, L; Levy, SM; Warren, JJ; Broffitt, B. (2011). Amoxicillin use during early childhood and fluorosis of later developing tooth zones. J Public Health Dent 71: 229-235. https://hero.epa.gov/reference/10482418/
Howard, BE; Phillips, J; Miller, K; Tandon, A; Mav, D; Shah, MR; Holmgren, S; Pelch, KE; Walker, V; Rooney, AA; Macleod, M; Shah, RR; Thayer, K. (2016). SWIFT-Review: a text-mining workbench for systematic review. Systematic Reviews 5: 87. https://hero.epa.gov/reference/4149688/
Howard, BE; Phillips, J; Tandon, A; Maharana, A; Elmore, R; Mav, D; Sedykh, A; Thayer, K; Merrick, BA; Walker, V; Rooney, A; Shah, RR. (2020). SWIFT-Active screener: accelerated document screening through active learning and integrated recall estimation. Environment International 138: 105623. https://hero.epa.gov/reference/6570105/
Huang, X; Chen, K; Wang, C; Gao, P. (2025). Characteristics of fluoride adsorption in different soil types: potential factors and implications for environmental risk assessment. Environmental Pollution 367: 125537. https://hero.epa.gov/reference/457543/
IARC (International Agency for Research on Cancer). (1987a). Fluorides. Environmental health criteria 227. https://hero.epa.gov/reference/13283245/
IARC (International Agency for Research on Cancer). (1987b). Summaries & evaluations. Fluorides (inorganic, used in drinking-water) (pp. 208). World Health Organization. https://hero.epa.gov/reference/13283246/
Idowu, OS; Azevedo, LB; Valentine, RA; Swan, J; Vasantavada, PV; Maguire, A; Zohoori, FV. (2019). The use of urinary fluoride excretion to facilitate monitoring fluoride intake: a systematic scoping review. PLoS ONE 14: e0222260. https://hero.epa.gov/reference/10482450/
Institute of Medicine. (1997a). Dietary reference intakes for calcium, phosphorus, magnesium, vitamin d, and fluoride. (ISBN 978-0-309-06350-0). Washington, DC: National Academy Press. https://hero.epa.gov/reference/1326493/
Institute of Medicine. (1997b). Dietary reference intakes for calcium, phosphorus, magnesium, vitamin d, and fluoride [Review]. The National Academies Collection: Reports funded by National Institutes of Health. https://hero.epa.gov/reference/13058559/
IPCS (International Programme on Chemical Safety). (2010). Characterization and application of physiologically based pharmacokinetic models in risk assessment. (Harmonization Project Document No 9). Geneva, Switzerland: World Health Organization. https://hero.epa.gov/reference/1064741/
Ishii, T; Suckling, G. (1991). The severity of dental fluorosis in children exposed to water with a high fluoride content for various periods of time. J Dent Res 70: 952-956. https://hero.epa.gov/reference/12158788/
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. https://hero.epa.gov/reference/50944/
Jean, KJ; Wassef, N; Gagnon, F; Valcke, M. (2018). A physiologically-based pharmacokinetic modeling approach using biomonitoring data in order to assess the contribution of drinking water for the achievement of an optimal fluoride dose for dental health in children. International Journal of Environmental Research and Public Health 15. https://hero.epa.gov/reference/10481731/

Kampouri, M; Gustin, K; Stråvik, M; Barman, M; Sandin, A; Sandberg, AS; Wold, AE; Vahter, M; Kippler, M. (2023). Associations of gestational and early-life exposure to toxic metals and fluoride with a diagnosis of food allergy or atopic eczema at 1 year of age. Environ Int 178: 108071. https://hero.epa.gov/reference/11811881/

Kerr-Wilson, CO; Mackay, DF; Smith, GCS; Pell, JP. (2012). Meta-analysis of the association between preterm delivery and intelligence. Journal of Public Health (Online) 34: 209-216. https://hero.epa.gov/reference/13283248/

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://hero.epa.gov/reference/10489160/

Kimmel, CA; Gaylor, DW. (1988). Issues in qualitative and quantitative risk analysis for developmental toxicology. Risk Analysis 8: 15-20. https://hero.epa.gov/reference/4250/

Kivrak, Y; Kars, T. (2012). Effects of fluoride on anxiety and depression in mice (pp. 302-306). https://hero.epa.gov/reference/1931650/

Koop, CE. (1984). Letter to William D. Ruckelshaus, Administrator, EPA, dated January 23, 1984, and transcript of proceedings of Surgeon General's Ad Hoc Committee on "Non-Dental Health Effects of Fluoride"; April 18-18, 1993, Jay R. Shapiro, Chairman. https://hero.epa.gov/reference/11264447/

Kraenzlin, ME; Kraenzlin, C; Farley, SM; Fitzsimmons, RJ; Baylink, DJ. (1990). Fluoride pharmacokinetics in good and poor responders to fluoride therapy. J Bone Miner Res 5 Suppl 1: S49-S52. https://hero.epa.gov/reference/13033768/

Kramer, RA; Allen, L; Gergen, PJ. (1995). Health and social characteristics and children's cognitive functioning: results from a national cohort. Am J Public Health 85: 312-318. https://hero.epa.gov/reference/13283209/

Krieg, EF; Chrislip, DW; Letz, RE; Otto, DA; Crespo, CJ; Brightwell, WS; Ehrenberg, RL. (2001). Neurobehavioral test performance in the third national health and nutrition examination survey. Neurotoxicology and Teratology 23: 569-589. https://hero.epa.gov/reference/6712044/

Kumar, JV; Swango, PA. (1999). Fluoride exposure and dental fluorosis in Newburgh and Kingston, New York: policy implications. Community Dent Oral Epidemiol 27: 171-180. https://hero.epa.gov/reference/12159161/

Lamar, W; Seegmiller, C. (1941). Determination of small quantities of fluoride in water, a modified zirconium-alizarin method. Ind Eng Chem Anal Ed 13: 901-902. https://hero.epa.gov/reference/13257281/

Lamar, WL. (1945). Determination of fluoride in water. a modified zirconium-alizarin method. Industrial and Engineering Chemistry Analytical Edition 148-149. https://hero.epa.gov/reference/13257285/

Lamar, WL; Drake, PG. (1955). Factors affecting the determination of fluoride in water with zirconium-alizarin. Journal of the American Water Works Association 563-572. https://hero.epa.gov/reference/13257280/

Lavalle-Carrasco, J; Molina-Frechero, N; Nevárez-Rascón, M; Sánchez-Pérez, L; Hamdan-Partida, A; González-González, R; Cassi, D; Isiordia-Espinoza, MA; Bologna-Molina, R. (2021). Recent biomarkers for monitoring the systemic fluoride levels in exposed populations: a systematic review. International Journal of Environmental Research and Public Health 18. https://hero.epa.gov/reference/10485686/

Lee, WS; Kim, JH; Han, B; Lee, GC; Jung, HR; Shin, YJ; Ha, EK; Lee, E; Lee, S; Han, MY. (2024). Association of fluoride exposure with disease burden and neurodevelopment outcomes in children in South Korea. World J Pediatr 20: 1029-1042.

https://hero.epa.gov/reference/12159057/
Leisenring, W; Ryan, L. (1992). Statistical properties of the NOAEL. Regul Toxicol Pharmacol 15: 161-171. https://hero.epa.gov/reference/84840/
Levy, SM; Kohout, FJ; Guha-Chowdhury, N; Kiritsy, MC; Heilman, JR; Wefel, JS. (1995). Infants' fluoride intake from drinking water alone, and from water added to formula, beverages, and food. J Dent Res 74: 1399-1407. https://hero.epa.gov/reference/1060467/
Li, M; Cui, J; Gao, Y; Zhang, W; Sun, L; Liu, X; Liu, Y; Sun, D. (2015). Pathological changes and effect on the learning and memory ability in rats exposed to fluoride and aluminum. Toxicology Research 4: 1366-1373. https://hero.epa.gov/reference/6579694/
Li, Y; Liang, C; Slemenda, CW; Ji, R; Sun, S; Cao, J; Emsley, CL; Ma, F; Wu, Y; Ying, P; Zhang, Y; Gao, S; Zhang, W; Katz, BP; Niu, S; Cao, S; Johnston, CC. (2001). Effect of long-term exposure to fluoride in drinking water on risks of bone fractures. J Bone Miner Res 16: 932-939. https://hero.epa.gov/reference/2528303/
Limón-Pacheco, JH; Jiménez-Córdova, MI; Cárdenas-González, M; Sánchez Retana, IM; Gonsebatt, ME; Del Razo, LM. (2018). Potential co-exposure to arsenic and fluoride and biomonitoring equivalents for Mexican children [Review]. Ann Glob Health 84: 257-273. https://hero.epa.gov/reference/5083828/
Liu, Y; Téllez-Rojo, M; Hu, H; Sánchez, BN; Martinez-Mier, EA; Basu, N; Mercado-García, A; Solano-González, M; Peterson, KE. (2019). Fluoride exposure and pubertal development in children living in Mexico City. Environ Health 18: 26. https://hero.epa.gov/reference/6405591/
Lu, Y; Luo, Q; Cui, H; Deng, H; Kuang, P; Liu, H; Fang, J; Zuo, Z; Deng, J; Li, Y; Wang, X; Zhao, L. (2017). Sodium fluoride causes oxidative stress and apoptosis in the mouse liver. Aging 9: 1623-1639. https://hero.epa.gov/reference/10481185/
Mascarenhas, AK. (2000). Risk factors for dental fluorosis: a review of the recent literature. Pediatr Dent 22: 269-277. https://hero.epa.gov/reference/13067887/
Mascarenhas, AK; Burt, BA. (1998). Fluorosis risk from early exposure to fluoride toothpaste. Community Dent Oral Epidemiol 26: 241-248. https://hero.epa.gov/reference/13028825/
Maurer, J; Cheng, M; Boysen, B; Anderson, R. (1990). 2-year carcinogenicity study of sodium-fluoride in rats. Journal of the National Cancer Institute 82: 1118-1126. https://hero.epa.gov/reference/12171098/
McClure, FJ. (1943). Ingestion of fluoride and dental caries: quantitative relations based on food and water requirements of children 1-12 years old. Am J Dis Child 66: 362-369. https://hero.epa.gov/reference/10510384/
McLanahan, ED; El-Masri, HA; Sweeney, LM; Kopylev, LY; Clewell, HJ; Wambaugh, JF; Schlosser, PM. (2012). Physiologically based pharmacokinetic model use in risk assessment--Why being published is not enough. Toxicological Sciences 126: 5-15. https://hero.epa.gov/reference/1015422/
Megregian, S; Maier, FJ. (1952). Modified zirconium-alizarin reagent for determination of fluoride in water. Journal of the American Water Works Association 239-246. https://hero.epa.gov/reference/13257282/
Morrow, CE; Culbertson, JL; Accornero, VH; Xue, L; Anthony, JC; Bandstra, ES. (2006). Learning disabilities and intellectual functioning in school-aged children with prenatal cocaine exposure. Dev Neuropsychol 30: 905-931. https://hero.epa.gov/reference/2856378/
NASEM (National Academies of Sciences, Engineering, and Medicine). (2017). Application of systematic review methods in an overall strategy for evaluating low-dose toxicity from endocrine active chemicals. In Consensus Study Report. Washington, D.C.: The National Academies Press. https://hero.epa.gov/reference/3982546/

NASEM (National Academies of Sciences, Engineering, and Medicine). (2021). Review of U.S. EPA's ORD staff handbook for developing IRIS assessments: 2020 version. (ISBN 9780309090889). Washington, DC: National Academies Press. https://hero.epa.gov/reference/9959764/

Nestorov, IA; Aarons, LJ; Arundel, PA; Rowland, M. (1998). Lumping of whole-body physiologically based pharmacokinetic models. J Pharmacokinet Biopharm 26: 21-46. https://hero.epa.gov/reference/819538/

NIEHS (National Institute of Environmental Health Sciences). (2022). NIEHS report on evaluating features and application of neurodevelopmental tests in epidemiological studies: NIEHS report 01. In NIEHS Report Series. Research Triangle Park (NC): National Institute of Environmental Health Sciences. https://hero.epa.gov/reference/12058548/

NIH (National Institutes of Health). (2025). Fluoride fact sheet for health professionals. 2025. https://hero.epa.gov/reference/13037052/

Niu, R; Han, H; Zhang, Y; Wang, J; Zhang, J; Yin, W; Yin, X; Sun, Z; Wang, J. (2016). Changes in liver antioxidant status of offspring mice induced by maternal fluoride exposure during gestation and lactation. Biol Trace Elem Res 172: 172-178. https://hero.epa.gov/reference/4927532/

NRC (National Research Council). (1993). Health effects of ingested fluoride. Washington, D.C.: National Academy Press. https://hero.epa.gov/reference/10489881/

NRC (National Research Council). (2006). Fluoride in drinking water: a scientific review of EPA's standards. Washington, DC: National Academies Press. https://hero.epa.gov/reference/11328274/

NRC (National Research Council). (2014). Review of EPA's integrated risk information system (IRIS) process. Washington, DC: The National Academies Press. https://hero.epa.gov/reference/2345577/

NTP (National Toxicology Program). (2016). NTP research report on systematic literature review on the effects of fluoride on learning and memory in animal studies: research report 1. In NTP Research Reports. Research Triangle Park (NC): National Toxicology Program. https://hero.epa.gov/reference/7266588/

NTP (National Toxicology Program). (2019). Handbook for conducting a literature-based health assessment using OHAT approach for systematic review and evidence integration. Research Triangle, NC: National Institute of Environmental Health Sciences. https://hero.epa.gov/reference/8803665/

NTP (National Toxicology Program). (2024). NTP monograph on the state of the science concerning fluoride exposure and neurodevelopment and cognition: a systematic review (pp. xviii - xix). Research Triangle Park, NC: U.S. Department of Health and Human Services. https://hero.epa.gov/reference/13037053/

OEHHA (Office of Environmental Health Hazard Assessment). (2008). TSD for noncancer RELs - appendix D.3 chronic RELs and toxicity summaries using the previous version of the hot spots risk assessment guidelines (OEHHA 1999). Sacramento, CA: California Office of Environmental Health Hazard Assessment: OEHHA. https://hero.epa.gov/reference/628645/

OEHHA (Office of Environmental Health Hazard Assessment). (2011). Evidence on the carcinogenicity of fluoride and its salts. https://hero.epa.gov/reference/13283243/

Parimi, N; Viswanath, V; Kashyap, B; Patil, PU. (2013). Hair as biomarker of fluoride exposure in a fluoride endemic area and a low fluoridated area. Int J Trichology 5: 148-150. https://hero.epa.gov/reference/12186652/

Pendrys, DG; Katz, RV; Morse, DE. (1996). Risk factors for enamel fluorosis in a nonfluoridated population. Am J Epidemiol 143: 808-815. https://hero.epa.gov/reference/13029841/

Peterson, E; Shapiro, H; Li, Y; Minnery, JG; Copes, R. (2016). Arsenic from community water fluoridation: quantifying the effect. J Water Health 14: 236-242.

https://hero.epa.gov/reference/3378868/
Picco, DC; Delbem, AC; Sassaki, KT; Sumida, DH; Antoniali, C. (2014). The effect of chronic treatment with fluoride on salivary activity, tooth, and bone in spontaneously hypertensive rats (SHR). Naunyn Schmiedebergs Arch Pharmacol 387: 321-328. https://hero.epa.gov/reference/13073081/
Plunkett, BA; Mele, L; Casey, BM; Varner, MW; Sorokin, Y; Reddy, UM; Wapner, RJ; Thorp, JM, Jr.; Saade, GR; Tita, ATN; Rouse, DJ; Sibai, B; Mercer, BM; Tolosa, JE; Caritis, SN. (2021). Association of breastfeeding and child IQ score at age 5 years. Obstet Gynecol 137: 561-570. https://hero.epa.gov/reference/11618904/
Pollick, HF. (2013). Salt fluoridation: a review [Review]. J Calif Dent Assoc 41: 395-397, 400-394. https://hero.epa.gov/reference/10488805/
Price, K; Haddad, S; Krishnan, K. (2003). Physiological modeling of age-specific changes in the pharmacokinetics of organic chemicals in children. Journal of Toxicology and Environmental Health, Part A: Current Issues 66: 417-433. https://hero.epa.gov/reference/819121/
Qian, M; Wang, D; Watkins, WE; Gebski, V; Yan, YQ; Li, M; Chen, ZP. (2005). The effects of iodine on intelligence in children: a meta-analysis of studies conducted in China. Asia Pac J Clin Nutr 14: 32-42. https://hero.epa.gov/reference/9601366/
Rango, T; Vengosh, A; Jeuland, M; Whitford, GM; Tekle-Haimanot, R. (2017). Biomarkers of chronic fluoride exposure in groundwater in a highly exposed population. Sci Total Environ 596-597: 1-11. https://hero.epa.gov/reference/7332431/
Rao, HV; Beliles, RP; Whitford, GM; Turner, CH. (1995). A physiologically based pharmacokinetic model for fluoride uptake by bone. Regulatory Toxicology and Pharmacology 22: 30-42. https://hero.epa.gov/reference/14160/
Reddy, PS; Pushpalatha, T; Reddy, PS. (2007). Suppression of male reproduction in rats after exposure to sodium fluoride during early stages of development. Naturwissenschaften 94: 607-611. https://hero.epa.gov/reference/1412648/
Richards, A; Mosekilde, L; Sogaard, CH. (1994). Normal age-related-changes in fluoride content of vertebral trabecular bone – relation to bone quality. Bone 15: 21-26. https://hero.epa.gov/reference/13029803/
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://hero.epa.gov/reference/10483230/
Rocha-Amador, D; Navarro, ME; Carrizales, L; Morales, R; Calderón, J. (2007). Decreased intelligence in children and exposure to fluoride and arsenic in drinking water. Cad Saude Publica 23 Suppl 4: S579-587. https://hero.epa.gov/reference/176992/
Ronfani, L; Vecchi Brumatti, L; Mariuz, M; Tognin, V; Bin, M; Ferluga, V; Knowles, A; Montico, M; Barbone, F. (2015). The complex interaction between home environment, socioeconomic status, maternal IQ and early child neurocognitive development: a multivariate analysis of data collected in a newborn cohort study. PLoS ONE 10: e0127052. https://hero.epa.gov/reference/13283252/
Rugg-Gunn, AJ; Rugg-Gunn, AJ; Enrique Villa, A; Rabelo Buzalaf, MA. (2011). Contemporary biological markers of exposure to fluoride [Review]. In Monographs in Oral Science. https://hero.epa.gov/reference/7386893/
Russ, TC; Killin, LOJ; Hannah, J; Batty, GD; Deary, IJ; Starr, JM. (2020). Aluminium and fluoride in drinking water in relation to later dementia risk. Br J Psychiatry 216: 29-34. https://hero.epa.gov/reference/7711368/
Saha, A; Mukherjee, AK; Ravichandran, B. (2016). Musculoskeletal problems and fluoride exposure: a cross-sectional study among metal smelting workers. Toxicol Ind Health 32: 1581-1588. https://hero.epa.gov/reference/3013945/

Sanchis, JM. (1934). Determination of fluorides in natural waters. Industrial and Engineering Chemistry Analytical Edition 134-135. https://hero.epa.gov/reference/13257284/

Sanchis, JM. (1936). Methods of testing and significance of fluorine and fluorides in water supplies. Journal of the American Water Works Association 1456-1468. https://hero.epa.gov/reference/13257278/

Sankhala, SS; Harshwal, R; Paliwal, P; Agarwal, A. (2014). Toe nails as a biomarker of chronic fluoride exposure secondary to high water fluoride content in areas with endemic fluorosis. Fluoride 47: 235-240. https://hero.epa.gov/reference/10481449/

Scott, RD. (1941). Modification of the fluoride determination. AWWA Water Sci 33: 2018-2020. https://hero.epa.gov/reference/13257277/

Shapiro, JR. (1983). Report to the surgeon general: by the ad hoc committee on the non-dental health effects of fluoride in drinking water. Ad Hoc Committee on the Non-Dental Health Effects of Fluoride in Drinking Water. https://hero.epa.gov/reference/11264440/

Sharashenidze, M; Tkeshelashvili, V; Nanobashvili, K. (2020). Dental fluorosis prevalence, severity and associated risk factors in pre-school aged children residing in fluoride deficient regions of Georgia. Georgian Med News 306: 57-61. https://hero.epa.gov/reference/10483042/

Shenkin, SD; Starr, JM; Deary, IJ. (2004). Birth weight and cognitive ability in childhood: a systematic review. Psychological Bulletin 130: 989-1013. https://hero.epa.gov/reference/13283247/

Skotowski, MC; Hunt, RJ; Levy, SM. (1995). Risk-factors for dental fluorosis in pediatric dental patients. J Public Health Dent 55: 154-159. https://hero.epa.gov/reference/13029827/

Soto-Barreras, U; Escalante-Villalobos, K; Holguín-Loya, B; Perez-Aguirre, B; Nevárez-Rascón, Z; MartÌnez-Martínez, R; Loyola-Rodríguez, J. (2019). Effect of fluoride in drinking water on dental caries and IQ in children (pp. 474-482). Chihuahua, Mexico. https://hero.epa.gov/reference/10481472/

Sowanou, A; Meng, X; Zhong, N; Ma, Y; Li, A; Wang, J; Li, H; Pei, J; Gao, Y. (2022). Association between osteoarthritis and water fluoride among tongyu residents, China, 2019: a case-control of population-based study. Biol Trace Elem Res 200: 3107-3116. https://hero.epa.gov/reference/10483053/

Spak, CJ; Ekstrand, J; Zylberstein, D. (1982). Bioavailability of fluoride added by baby formula and milk. Caries Res 16: 249-256. https://hero.epa.gov/reference/13033443/

Streissguth, An; Barr, HM; Sampson, PD; Darby, BL; Martin, DC. (1989). IQ at age 4 in relation to maternal alcohol use and smoking during pregnancy. Developmental Psychology 3-11. https://hero.epa.gov/reference/13283250/

Streissguth, AP; Barr, HM; Sampson, PD. (1990). Moderate prenatal alcohol exposure: effects on child IQ and learning problems at age 7 1/2 years. Alcoholism, clinical and experimental research 14: 662-669. https://hero.epa.gov/reference/13283249/

Sun, Z; Li, S; Guo, Z; Li, R; Wang, J; Niu, R; Wang, J. (2018). Effects of fluoride on SOD and CAT in testis and epididymis of mice. Biol Trace Elem Res 184: 148-153. https://hero.epa.gov/reference/10483211/

Taher, MK; Momoli, F; Go, J; Hagiwara, S; Ramoju, S; Hu, X; Jensen, N; Terrell, R; Hemmerich, A; Krewski, D. (2024). Systematic review of epidemiological and toxicological evidence on health effects of fluoride in drinking water. Critical Reviews in Toxicology 54: 2-34. https://hero.epa.gov/reference/12344252/

Tan, YM; Chan, M; Chukwudebe, A; Domoradzki, J; Fisher, J; Hack, CE; Hinderliter, P; Hirasawa, K; Leonard, J; Lumen, A; Paini, A; Qian, H; Ruiz, P; Wambaugh, J; Zhang, F; Embry, M. (2020). PBPK model reporting template for chemical risk assessment applications [Review]. Regulatory Toxicology and Pharmacology 115: 104691. https://hero.epa.gov/reference/6592021/

Taras, MJ; Cisco, HD; Garnell, M. (1950). Interferences in alizarin method of fluoride determination. Journal of the American Water Works Association 583-588. https://hero.epa.gov/reference/13257279/

Taylor, KW; Eftim, SE; Sibrizzi, CA; Blain, RB; Magnuson, K; Hartman, PA; Rooney, AA; Bucher, JR. (2025). Fluoride exposure and children's IQ scores: a systematic review and meta-analysis [Comment]. JAMA Pediatrics 179: 282-292. https://hero.epa.gov/reference/12343370/

Thomas, DB; Basu, N; Martinez-Mier, EA; Sánchez, BN; Zhang, Z; Liu, Y; Parajuli, RP; Peterson, K; Mercado-García, A; Bashash, M; Hernández-Avila, M; Hu, H; Téllez-Rojo, MM. (2016). Urinary and plasma fluoride levels in pregnant women from Mexico City. Environ Res 150: 489-495. https://hero.epa.gov/reference/6703062/

Thornton-Evans, G; Junger, ML; Lin, M; Wei, L; Espinoza, L; Beltran-Aguilar, E. (2019). Use of toothpaste and toothbrushing patterns among children and adolescents - United States, 2013-2016. MMWR Morb Mortal Wkly Rep 68: 87-90. https://hero.epa.gov/reference/10488734/

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://hero.epa.gov/reference/6718311/

Trasande, L; Landrigan, PJ; Schechter, C. (2005). Public health and economic consequences of methyl mercury toxicity to the developing brain. Environ Health Perspect 113: 590-596. https://hero.epa.gov/reference/528912/

Turkheimer, E; Haley, A; Waldron, M; D'Onofrio, B; Gottesman, II. (2003). Socioeconomic status modifies heritability of IQ in young children. Psychological Science 14: 623-628. https://hero.epa.gov/reference/13283210/

U.S. Department of Health and Human Services. (2015). U.S. Public health service recommendation for fluoride concentration in drinking water for the prevention of dental caries. Public Health Rep 130: 318-331. https://hero.epa.gov/reference/10490312/

U.S. EPA (U.S. Environmental Protection Agency). (1986a). Guidelines for carcinogenic risk assessment (pp. 33992-34003). (Federal Register 51(185)). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/13036530/

U.S. EPA (U.S. Environmental Protection Agency). (1986b). National primary and secondary drinking water regulations; fluoride; final rule. Fed Regist 51: 11395-11412. https://hero.epa.gov/reference/10634795/

U.S. EPA (U.S. Environmental Protection Agency). (1987). Fluorine (soluble fluoride); CASRN 7782-41-4. integrated risk information system (IRIS) chemical assessment summary. Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/10186180/

U.S. EPA (U.S. Environmental Protection Agency). (1988). Recommendations for and documentation of biological values for use in risk assessment [EPA Report]. (EPA600687008). Cincinnati, OH: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/64560/

U.S. EPA (U.S. Environmental Protection Agency). (1991). Guidelines for developmental toxicity risk assessment. Fed Regist 56: 63798-63826. https://hero.epa.gov/reference/732120/

U.S. EPA (U.S. Environmental Protection Agency). (1994). Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry [EPA Report]. (EPA600890066F). Research Triangle Park, NC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/6488/

U.S. EPA (U.S. Environmental Protection Agency). (1995). The use of the benchmark dose approach in health risk assessment. (EPA/630/R-94/007). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://hero.epa.gov/reference/5992/

U.S. EPA (U.S. Environmental Protection Agency). (1996a). Guidelines for reproductive toxicity risk

assessment [EPA Report]. Fed Regist 61: 56274-56322. https://hero.epa.gov/reference/7581/
U.S. EPA (U.S. Environmental Protection Agency). (1996b). Reregistration eligibility decision (RED): cryolite. (includes RED facts: cryolite fact sheet). (NTIS/02992829). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/2072993/
U.S. EPA (U.S. Environmental Protection Agency). (1998). Guidelines for neurotoxicity risk assessment. (EPA/630/R-95/001F). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/30021/
U.S. EPA (U.S. Environmental Protection Agency). (2000). Policy and program requirements for the mandatory agency-wide quality system. (CIO 2105.0 (formerly 5360.1 A2)). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/1777862/
U.S. EPA (U.S. Environmental Protection Agency). (2002a). Guidelines for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated by the Environmental Protection Agency [EPA Report]. (EPA/260R-02-008). Washington, DC: U.S. Environmental Protection Agency, Office of Environmental Information. https://hero.epa.gov/reference/635281/
U.S. EPA (U.S. Environmental Protection Agency). (2002b). A review of the reference dose and reference concentration processes. (EPA630P02002F). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/88824/
U.S. EPA (U.S. Environmental Protection Agency). (2003). Six-year review 1 of drinking water standards [Website]. https://hero.epa.gov/reference/10368502/
U.S. EPA (U.S. Environmental Protection Agency). (2004). Toxicological review of boron and compounds. in support of summary information on the integrated risk information system (IRIS) [EPA Report]. (EPA/635/04/052). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/198783/
U.S. EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment [EPA Report]. (EPA630P03001F). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/6324329/
U.S. EPA (U.S. Environmental Protection Agency). (2006). Approaches for the application of physiologically based pharmacokinetic (PBPK) models and supporting data in risk assessment (final report) [EPA Report] (pp. 1-123). (EPA/600/R-05/043F). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment. https://hero.epa.gov/reference/194568/
U.S. EPA (U.S. Environmental Protection Agency). (2007). Reregistration eligibility decision (RED) for sodium fluoride (pp. 71). (EPA 739-R-07-010). U.S. Environmental Protection Agency, Office of Prevention, Pesticides, and Toxic Substances. https://hero.epa.gov/reference/13037045/
U.S. EPA (U.S. Environmental Protection Agency). (2008a). EPA Procedure for Quality Policy. (EPA Classification No.: CIO 2106-P-01.0). Office of Environmental Information. https://hero.epa.gov/reference/1777870/
U.S. EPA (U.S. Environmental Protection Agency). (2008b). Fluoride: dose-response analysis for non-cancer effects. Dental fluorosis: evaluations of key studies [EPA Report]. (EPA 820-R-10-019). Washington, DC: U.S. Environmental Protection Agency, Office of Water. https://hero.epa.gov/reference/8151256/
U.S. EPA (U.S. Environmental Protection Agency). (2008c). Fluoride: dose-response analysis for non-cancer effects. Fluoride-related skeletal effects: evaluation of key studies. https://hero.epa.gov/reference/13283244/
U.S. EPA (U.S. Environmental Protection Agency). (2010a). Fluoride: dose-response analysis for non-cancer effects [EPA Report]. (EPA 820-R-10-019). Washington, DC: U.S. Environmental Protection Agency, Office of Water. https://hero.epa.gov/reference/10493692/

U.S. EPA (U.S. Environmental Protection Agency). (2010b). Fluoride: exposure and relative source contribution analysis. (820R10015). Washington, DC. https://hero.epa.gov/reference/10493698/

U.S. EPA (U.S. Environmental Protection Agency). (2011a). Exposure factors handbook 2011 edition (final report). Chapter 3 — Ingestion of water and other select liquids [EPA Report]. (EPA/600/R-09/052F). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development. https://hero.epa.gov/reference/10721945/

U.S. EPA (U.S. Environmental Protection Agency). (2011b). Recommended use of body weight 3/4 as the default method in derivation of the oral reference dose. (EPA100R110001). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/752972/

U.S. EPA (U.S. Environmental Protection Agency). (2011c). Toxicological review of trichloroethylene (CAS No. 79-01-6) in support of summary information on the integrated risk information system (IRIS) [EPA Report]. (EPA635R09011F). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/3532116/

U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance [EPA Report]. (EPA100R12001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://hero.epa.gov/reference/1239433/

U.S. EPA (U.S. Environmental Protection Agency). (2014a). Framework for human health risk assessment to inform decision making. final [EPA Report]. (EPA/100/R-14/001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://hero.epa.gov/reference/2324779/

U.S. EPA (U.S. Environmental Protection Agency). (2014b). Guidance for applying quantitative data to develop data-derived extrapolation factors for interspecies and intraspecies extrapolation [EPA Report]. (EPA/100/R-14/002F). Washington, DC: Risk Assessment Forum, Office of the Science Advisor. https://hero.epa.gov/reference/2520260/

U.S. EPA (U.S. Environmental Protection Agency). (2016). Chemical contaminant summaries for the third six-year review of existing national primary drinking water regulations [EPA Report]. Washington, DC: U.S. Environmental Protection Agency, Office of Water. https://hero.epa.gov/reference/10494328/

U.S. EPA (U.S. Environmental Protection Agency). (2017). National primary drinking water regulations; announcement of the results of EPA's review of existing drinking water standards and request for public comment and/or information on related issues (pp. 3518-3552). (EPA–HQ–OW–2016–0627; FRL–9957–49– OW). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/13283203/

U.S. EPA (U.S. Environmental Protection Agency). (2019a). Exposure factors handbook chapter 3 (update): ingestion of water and other select liquids [EPA Report]. (EPA/600/R-18/259F). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/7267482/

U.S. EPA (U.S. Environmental Protection Agency). (2019b). Systematic review protocol for the hexavalent chromium IRIS assessment (preliminary assessment materials) [EPA Report]. (EPA/635/R-18/155). Environmental Protection Agency. https://hero.epa.gov/reference/5097854/

U.S. EPA (U.S. Environmental Protection Agency). (2019c). Systematic review protocol for the polychlorinated biphenyls (PCBs) noncancer IRIS assessment (preliminary assessment materials) [EPA Report]. (EPA/635/R-19/201). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment, Integrated Risk Information System. https://hero.epa.gov/reference/7415523/

U.S. EPA (U.S. Environmental Protection Agency). (2019d). Title XIV of the public health service act safety of public water systems (safe drinking water act) as amended through Pub. L. No. 116-92.

https://hero.epa.gov/reference/5935484/
U.S. EPA (U.S. Environmental Protection Agency). (2020). Systematic review protocol for the methylmercury (MeHg) IRIS assessment (preliminary assessment materials). (EPA/635/R-19/243). Washington, DC. https://hero.epa.gov/reference/7607843/
U.S. EPA (U.S. Environmental Protection Agency). (2022a). IRIS toxicological review of perfluorobutanoic acid and related salts [EPA Report]. (EPA/635/R-22/277Fa). Washington, D.C.: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/10668899/
U.S. EPA (U.S. Environmental Protection Agency). (2022b). ORD staff handbook for developing IRIS assessments. (EPA600R22268). Washington, DC: U.S. Environmental Protection Agency, Office of Research and Development, Center for Public Health and Environmental Assessment. https://hero.epa.gov/reference/10367891/
U.S. EPA (U.S. Environmental Protection Agency). (2023a). EPA protocol for the fourth review of existing national primary drinking water regulations. (EPA 815-R-24-018). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/11346388/
U.S. EPA (U.S. Environmental Protection Agency). (2023b). ORD human health toxicity value for lithium bis [(trifluoromethyl)sulfonyl]azanide (HQ-115). (EPA/600/R-22/195F). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/11311313/
U.S. EPA (U.S. Environmental Protection Agency). (2023c). A systematic evidence map of noncancer health endpoints and exposures to polychlorinated biphenyl (PCB) mixtures. In EPA IRIS Assessments. (EPA/600/R-22/201). Washington, DC: Environmental Protection Agency. https://hero.epa.gov/reference/10443555/
U.S. EPA (U.S. Environmental Protection Agency). (2023d). Toxicological review of perfluorohexanoic acid (PFHxA) and related salts (Final Report, 2023) [EPA Report]. (EPA/635/R-23/027F). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/11133619/
U.S. EPA (U.S. Environmental Protection Agency). (2024a). Final human health toxicity assessment for perfluorooctane sulfonic acid (PFOS) and related salts [EPA Report]. Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://hero.epa.gov/reference/11350937/
U.S. EPA (U.S. Environmental Protection Agency). (2024b). Final human health toxicity assessment for perfluorooctanoic acid (PFOA) and related salts [EPA Report]. Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://hero.epa.gov/reference/11350934/
U.S. EPA (U.S. Environmental Protection Agency). (2024c). Final. appendix: human health toxicity assessment for perfluorooctanoic acid (PFOA) and related salts. (815R24008). Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://hero.epa.gov/reference/11414343/
U.S. EPA (U.S. Environmental Protection Agency). (2024d). Integrated science assessment (ISA) for lead (final report). (EPA/600/R-23/375). Washington, DC: U.S. Environnmental Protection Agency (U.S. EPA). https://hero.epa.gov/reference/11854650/
U.S. EPA (U.S. Environmental Protection Agency). (2024e). IRIS toxicological review of hexavalent chromium [Cr(VI)]. (EPA/635/R-24/164Fa). Washington, DC: U.S. Environnmental Protection Agency (U.S. EPA). https://hero.epa.gov/reference/11854652/
U.S. EPA (U.S. Environmental Protection Agency). (2024f). Results of the Health Effects Assessment for the Fourth Six-Year Review of Existing Chemical and Radionuclide National Primary Drinking Water Standards. (EPA-815-R-24-020). Washington, D.C.: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://hero.epa.gov/reference/13037046/

U.S. EPA (U.S. Environmental Protection Agency). (2024g). Review of EPA's draft IRIS toxicological review of inorganic arsenic: final report. (EPA-SAB-25-004). Washington, D.C.: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/12106843/

U.S. EPA (U.S. Environmental Protection Agency). (2024h). Toxicological review of perfluorodecanoic acid (PFDA) and related salts (final report, 2024). (EPA/635/R-24/172Fa). Washington, DC: U.S. Environmental Protection Agency. https://hero.epa.gov/reference/11830815/

U.S. EPA (U.S. Environmental Protection Agency). (2025a). Draft health effects technical support document: deriving a reference dose and maximum contaminant level goal for perchlorate in drinking water. Washington, DC: U.S. Environmental Protection Agency, Office of Water, Health and Ecological Criteria Division. https://hero.epa.gov/reference/12338752/

U.S. EPA (U.S. Environmental Protection Agency). (2025b). EPA will expeditiously review new science on fluoride in drinking water. https://hero.epa.gov/reference/13283238/

U.S. EPA (U.S. Environmental Protection Agency). (2025c). IRIS toxicological review of inorganic arsenic. U.S. Environmental Protection Agency. https://hero.epa.gov/reference/12232885/

U.S. EPA (U.S. Environmental Protection Agency). (2025d). Toxicological review of perfluorohexanesulfonic acid (PFHxS) and related salts: final report, 2025 [EPA Report]. (EPA/635/R-25/012). U.S. Environmental Protection Agency. https://hero.epa.gov/reference/12336103/

U.S. EPA (U.S. Environmental Protection Agency). (2025e). Updated problem formulation and protocol for the inorganic arsenic IRIS assessment. U.S. Environmental Protection Agency. https://hero.epa.gov/reference/13037050/

U.S. EPA (U.S. Environmental Protection Agency). (2026). Fluoride [CASRN 7782-41-4] human health toxicity assessment: preliminary assessment plan and literature survey. (EPA-821-P-26-001). Washington, DC: U.S. Environmental Protection Agency; Office of Ground Water and Drinking Water; Drinking Water Science and Engineering Division. https://hero.epa.gov/reference/13257200/

Underwood, EJ. (1977). Trace elements in human and animal nutrition (4 ed.). New York, NY: Academic Press. https://hero.epa.gov/reference/8528/

United States Executive Office of the President Donald Trump. (2025). Executive order 14303—Restoring gold standard science. Compilation of Presidential Documents. https://hero.epa.gov/reference/13283237/

Valdez Jiménez, L; López Guzmán, OD; Cervantes Flores, M; Costilla-Salazar, R; Calderón Hernández, J; Alcaraz Contreras, Y; Rocha-Amador, DO. (2017). In utero exposure to fluoride and cognitive development delay in infants. NeuroToxicology 59: 65-70. https://hero.epa.gov/reference/10482786/

Varghese, A; Cawley, M; Hong, T. (2018). Supervised clustering for automated document classification and prioritization: a case study using toxicological abstracts. Environ Syst Decis 38: 398-414. https://hero.epa.gov/reference/4449738/

Vieira, APG; Mousny, M; Maia, R; Hancock, R; Everett, ET; Grynpas, MD. (2005). Assessment of teeth as biomarkers for skeletal fluoride exposure. Osteoporos Int 16: 1576-1582. https://hero.epa.gov/reference/12171010/

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: 60-68. https://hero.epa.gov/reference/10488665/

Walker, OJ; Finlay, GR. (1940). The determination of small amounts of fluorine in water. Canadian Journal of Research 151-159. https://hero.epa.gov/reference/13257283/

Waterhouse, C; Taves, D; Munzer, A. (1980). Serum inorganic fluoride: changes related to previous

fluoride intake, renal function and bone resorption. Clin Sci (Lond) 58: 145-152. https://hero.epa.gov/reference/13033780/

Wheeler, M; Bailer, AJ. (2009). Comparing model averaging with other model selection strategies for benchmark dose estimation. Environmental and Ecological Statistics 16: 37-51. https://hero.epa.gov/reference/2316666/

Wheeler, MW; Davis, JA; Blessinger, T; Shao, K; Allen, BC; Olszyk, L; Gift, JS. (2019). Quantitative risk assessment: developing a Bayesian approach to dichotomous dose-response model averaging (Draft). https://hero.epa.gov/reference/5017804/

White, RF; Campbell, R; Echeverria, D; Knox, SS; Janulewicz, P. (2009). Assessment of neuropsychological trajectories in longitudinal population-based studies of children. Journal of Epidemiology and Community Health 63: i15-i26. https://hero.epa.gov/reference/534932/

WHO (World Health Organization). (1984). Guidelines for drinking-water quality: volume 1: recommendations (1st ed.). Geneva, Switzerland: World Health Organization. https://hero.epa.gov/reference/10606114/

WHO (World Health Organization). (2004). Fluoride in drinking-water. Background document for development of WHO guidelines for drinking-water quality. (WHO/SDE/WSH/03.04/96). Geneva. https://hero.epa.gov/reference/10634794/

Yeung, CA. (2014). Some beneficial effect on root caries from use of higher concentration fluoride toothpaste (5000 ppm F) [Comment]. Evid Based Dent 15: 8-9. https://hero.epa.gov/reference/10488640/

Zhang, C; McElroy, AC; Liberatore, HK; Alexander, NLM; Knappe, DRU. (2021). Stability of per- and polyfluoroalkyl substances in solvents relevant to environmental and toxicological analysis. Environmental Science & Technology 56: 6103-6112. https://hero.epa.gov/reference/9962000/

Zohoori, FV; Buzalaf, MA; Cardoso, CA; Olympio, KP; Levy, FM; Grizzo, LT; Mangueira, DF; Sampaio, FC; Maguire, A. (2013). Total fluoride intake and excretion in children up to 4 years of age living in fluoridated and non-fluoridated areas. Eur J Oral Sci 121: 457-464. https://hero.epa.gov/reference/10481220/

Zohoori, FV; Omid, N; Sanderson, RA; Valentine, RA; Maguire, A. (2019). Fluoride retention in infants living in fluoridated and non-fluoridated areas: effects of weaning. British Journal of Nutrition 121: 74-81. https://hero.epa.gov/reference/5421195/

Appendix A. Response to Public Comments on the Preliminary Assessment Plan and Literature Survey

A.1. Introduction

Section Overview: This section summarizes the public comments received by the EPA on the Preliminary Assessment Plan and Literature Survey and provides responses from the EPA. This document, the protocol, is an expanded version of the assessment plan that contains material from the assessment plan revised in response to public comments received, as well as new material specific to the protocol. In some cases, the new material specific to the protocol specifically addresses comments received from the public. In other cases, it provides additional detail on how the EPA will conduct the Fluoride Human Health Toxicity Assessment. The draft assessment will be made available for public comment and will be peer-reviewed.

The first planning step in developing the EPA's updated Fluoride Human Health Toxicity Assessment (*toxicity assessment* herein) was publishing the Preliminary Assessment Plan and Literature Survey (*assessment plan* herein) (U.S. EPA, 2026). The assessment plan described the EPA's preliminary plan for developing the toxicity assessment, including the intended scope or focus, areas of scientific complexity, and objectives, as well as methods for identifying relevant literature and the results of the initial literature search. The assessment plan was released for public comment period on January 28, 2026, and the comment period ended on February 27, 2026. The EPA received approximately 1,000 public comments, of which 521 were unique comments. The vast majority of the unique and substantive comments were generally supportive of the EPA's efforts to develop an updated fluoride human health toxicity assessment.

This Protocol for the Fluoride Human Health Toxicity Assessment (protocol herein) is an expanded version of the assessment plan and was informed by the public comments that were received. The protocol contains content in the assessment plan updated in response to public comments (Sections 1–4). The protocol also includes additional content on systematic review and analysis methods (Sections 5–8) that will be used to develop the toxicity assessment.

To prepare this Response to Comments document for the assessment plan, the agency summarized comments into 17 topic categories and developed responses to the summarized comments (see Sections A.3–A.5 below). The full text of all comments is available in the docket for this action (EPA-HQ-OW-2025-3823 at www.regulations.gov).

A.1.1. Overview of Public Comments and Impact on Protocol Development

In response to public comments, EPA revised the assessment plan content within this protocol, including (but not limited to) clarifications related to:

- Use of existing assessments developed by national and international health agencies
- Adversity of dental fluorosis
- Role of AI in identifying relevant studies

- Consideration of additional peer-reviewed publications identified by public comments

Many public comments focused on considerations that are relevant to assessment development steps beyond the scope of the assessment plan. Though these comments were not relevant to the assessment plan itself, in developing the protocol, which is intended to provide detailed information on methods, the EPA ensured that specific feedback provided by public commenters was taken into account. The methods topics included:

- Fluoride- and outcome-specific criteria for study evaluation of epidemiology studies (e.g., consideration of different measures to determine exposure levels, fluorosis grading, or neuropsychological scores and identification of important potential confounders)
- Methods for pharmacokinetic analysis, including criteria for evaluating existing pharmacokinetic (PK) models
- Consideration of susceptible populations
- Selection of studies for dose-response analysis
- Methods for dose-response analysis, including benchmark dose modeling
- Further detail on literature search and screening methods
- Plan for literature search update prior to toxicity assessment finalization

A.2. List of Commenters

Academy of General Dentistry (AGD)

American Academy of Cariology

American Academy of Pediatric Dentistry (AAPD)

American Academy of Pediatrics (AAP)

American Association for Dental, Oral, and Craniofacial Research (AADOCR)

American Association of Public Health Dentistry (AAPHD)

American Dental Association (ADA)

American Dental Education Association (ADEA)

American Dental Hygienists' Association (ADHA)

American Dental Therapy Association

American Fluoridation Society, Inc.

American Network of Oral Health Coalitions

American Network of Oral Health Coalitions (ANOHC)

American Public Health Association (APHA)

American Student Dental Association (ASDA)

American Water Works Association (AWWA) and Association of Metropolitan Water Agencies (AMWA)

Aqua General Inc.

Arcora Foundation

Arizona Academy of Pediatric Dentistry

Arizona Oral Health Coalition

Aroostook Dental Clinic, Inc./DBA St. Apollonia Dental Clinic

Association for Utah Community Health (AUCH)

Association of California Water Agencies (ACWA)

Association of Critical Access Dental Providers of Minnesota

Australian Dental Association (ADA)

Balster and Pfohl Pediatric Dentistry

Big Cities Health Coalition (BCHC)

Black Hills Pediatric Dentistry

Bluff City Dental

California Dental Association (CDA)
California Department of Public Health
California Society of Pediatric Dentistry (CSPD)
CareQuest Institute for Oral Health
Cedar Brook Pediatric Dentistry
Children's Health Alliance of Wisconsin
Citizens for Honesty
City of New York
Clackamas Health Centers Dental Program
Colorado Springs Utilities
Connecticut Oral Health Initiative (COHI)
Connecticut Society of Pediatric Dentists
Connecticut Society of Pediatric Dentists
Connecticut State Dental Association (CSDA)
Consumer Healthcare Products Association (CHPA)
Dunes Dental 4 Kids Dentists
First 5 Sacramento Commission
Florida Academy of Pediatric Dentistry
Fluoride Action Network (FAN)
Geisinger
Illinois Oral Health Coalition (IOHC)
International Association of Paediatric Dentistry
Iowa Academy of Pediatric Dentistry
Kentucky Voices for Health
King County Citizens Against Fluoridation (KCCAF).
Lakeside Smiles Pediatric Dentistry
Maine Primary Care Association (MPCA)
Maine Society of Pediatric Dentistry
Maricopa County Medical Society
Marquette County Health Department (MCHD)
Maryland Academy of Pediatric Dentistry (MdAPD)
Maryland State Dental Association
Medi-Cal Dental Advisory Committee Oral Health Equity Workgroup (MCDAC)
Michigan Dental Hygienists' Association (MDHA)
Mississippi Academy of Pediatric Dentistry (MAPD)
Mississippi Chapter of the American Academy of Pediatrics
Mississippi Dental Association
Mississippi State Medical Association (MSMA)
Missouri Dental Association (MDA)
National Consumers League (NCL)
National Network for Oral Health Access (NNOHA)
Nebraska Society of Pediatric Dentistry
Nest Pediatric Dentistry
New York State Association of County Health Officials (NYSACHO)
Norfolk Children's Dental Clinic, LLC
Northern Kentucky Water District (NKWD)
Northern Virginia Dental Society (VDA)
Ohio Academy of Pediatric Dentistry (OAPD)
Pennsylvania Academy of Pediatric Dentistry (PAPD)
Pennsylvania Coalition for Oral Health (PCOH)
Pennsylvania Dental Association (PDA)
PFPC Canada
Population Oral Health at UQ (POHUQ)
Precision Dental
Rhode Island Dental Association
Smiles for Life
Somerset Pediatric Dental Associates, LLC
South Dakota Society of Pediatric Dentistry
St. Clair County Health Department

The Dental Trade Alliance

The Rural Community Assistance Partnership (RCAP)

University of Washington School of Dentistry and School of Public Health

Utah Academy of Pediatric Dentistry

Utah Dental Hygienists' Association

Utah Oral Health Coalition

Utah Oral Health Coalition

Vermont Environmental Justice Network

Viera Pediatric Dentistry

Viera Pediatric Dentistry (VPD)

Virginia Academy of Pediatric Dentistry

Wash Park Pediatric Dentistry

Washington Action for Safe Water

Washington State Academy of Pediatric Dentistry

West Virginia Board of Dentistry

West Virginia Dental Hygienists' Association (WVDHA)

Wisconsin Academy of Pediatric Dentistry

Wisconsin Chapter of the American Academy of Pediatrics (WIAAP)

Yale New Haven Hospital

Individual and Anonymous Commenters

A.3. Proposed Scope of the Assessment

A.3.1. Conclusion That Neurodevelopment Is a Well-Established Hazard

Comment 3.1.1: Some commenters suggested that it was not appropriate for the EPA to conclude that neurodevelopment has been categorized as a 'well-established' health hazard before a literature review and synthesis is conducted.

EPA Response: The EPA disagrees that it was not appropriate to conclude that neurodevelopment is a well-established health hazard because the agency critically reviewed multiple recent health effects documents from national and international health agencies (EFSA Scientific Committee, 2025; Health Canada, 2024; NTP, 2024; Taher et al., 2024) that reviewed the health effects literature and identified consistency in hazard conclusions for neurodevelopmental studies. These agencies used Gold Standard Science including systematic literature searches and evaluation of risk of bias to determine that neurodevelopmental effects are associated with fluoride exposure. These conclusions were based on a large body of evidence including numerous studies judged as low risk of bias. The EPA's critical review of these thorough investigations from reliable sources led the agency to conclude that neurodevelopmental toxicity is a hazard of fluoride. This means that evidence supports the conclusion that fluoride exposure can result in neurodevelopmental effects, such as impaired cognition/reduced IQ. This does not imply that the adverse effects can be observed at all dose/exposure levels. Subsequent dose-response assessment, which follows hazard identification, will be conducted as part of the Fluoride Human Health Toxicity Assessment to characterize the dose or exposure below which adverse effects are not anticipated. The EPA has a precedent for accepting well-established hazards (e.g., recent protocols of methyl mercury (U.S. EPA, 2020) and hexavalent chromium (U.S. EPA, 2019b)) and this approach was encouraged, in general, by the National Academies of Science, Engineering, and Medicine (NASEM, 2017). Accepting well-established hazards after critical evaluation of the available reports is consistent with Gold Standard Science objectives and allows the agency to avoid duplication of effort thereby saving time to complete the assessment and focusing resources on determining the currently unknown level at which these neurodevelopmental effects occur (i.e., dose-response analysis).

Comment 3.1.2: Some commenters indicated that there is "not a sufficient basis" to set a reference dose (RfD) based on neurodevelopment. For example, one commenter said, "[t]he National Toxicology Program (NTP) report described "moderate" confidence that high fluoride levels are "consistently associated" with lower intelligence quotient (IQ) in children. Said differently, the report conveys that there is enough information to warrant further study, but not a sufficient basis to set an RfD based on this health endpoint." Similarly, one commenter stated that due to variability in outcome and exposure measurement, the body of neurodevelopmental findings "cannot serve as a dependable basis for national policy."

EPA Response: The EPA disagrees with commenters' assertions that there is not a sufficient basis to establish an RfD for fluoride based on neurodevelopment, and as explained further below, disagrees with the underlying premise of this assertion. While the commenters did not

describe the level of evidence considered "sufficient basis" for deriving an RfD, the EPA is following agency human health risk assessment guidance and methodology and Gold Standard Science principles to make the determination about sufficient basis for establishing hazard. Moderate evidence in human studies is commonly used for toxicity value derivation. As described in the EPA's peer-reviewed methodology for developing toxicity values (U.S. EPA, 2022b), dose-response analysis is generally performed for health effect categories with a summary evidence integration judgment of "evidence demonstrates" or "evidence indicates (likely)," the latter of which can be supported by "moderate" human evidence. Whether the available data are suitable for dose-response analysis will be evaluated during assessment development. Regarding the comment about the NTP report, the EPA disagrees that the report concluded that there "not a sufficient basis to set an RfD based on this health endpoint." The purpose of the NTP monograph was to determine whether there was an association between fluoride exposure and neurodevelopmental outcomes, not to derive an RfD; toxicity values are beyond the scope of the NTP monograph. The NTP monograph did not make conclusions about data sufficiency for the purpose of deriving an RfD. The EPA will independently review available data to determine whether there is a sufficient basis to set an RfD for adverse neurodevelopmental effects. Methods for study evaluation are described in Section 6, and considerations for suitability for dose-response are described in Section 8.

A.3.2. Comments Regarding the National Toxicology Program (NTP) Report on Fluoride and Neurodevelopment

Comment 3.2.1: Some commenters stated that the NTP monograph and corresponding meta-analysis "features several methodological flaws" and noted that "NASEM reviews identified shortcomings" in the drafts of the NTP monograph. Of note, commenters stated that there were issues with "reproducibility and transparency of the systematic review process" and with "risk of bias analysis," including "insufficient and/or inconsistent evaluation of confounding" (e.g., not having iodine as a key confounder) and "lack of rigorous statistical review."

EPA Response: The EPA disagrees with the commenters' characterization of the NTP monograph and corresponding meta-analysis. The purpose of the NTP monograph was to describe the state of the science concerning fluoride exposure and neurodevelopment and cognition. The NTP monograph was developed using rigorous systematic review methods (Office of Health Assessment and Translation (OHAT) methodology) that have been widely accepted in the environmental health community. While some peer reviewers and commenters scientifically disagreed regarding the specific decisions made during the development of the report, peer reviewers agreed with the overall conclusion of moderate confidence "that higher levels of fluoride exposure, such as drinking water containing more than 1.5 milligrams of fluoride per liter, are associated with lower IQ in children" (NTP, 2024). Further, this conclusion is consistent with the findings of other authoritative bodies as discussed in Section A.3.1. Since the EPA is conducting an independent review of the literature to identify studies suitable for dose-response, any decisions regarding specific studies made by NTP will not impact the EPA's conclusions in the Fluoride Human Health Toxicity Assessment. Furthermore, the EPA has transparently described the study evaluation criteria (Section 6.2) that will be applied in the

toxicity assessment and the implementation of those criteria in the draft assessment will be subject to external peer review.

Comment 3.2.2: Related to the studies comprising the NTP monograph, commenters noted that "approximately three fourths of the studies reviewed were of low quality and high risk of bias."

EPA Response: Consideration of bias is an important part of any systematic review of the scientific literature. Risk of bias in a study is an indication of how well the study was conducted and the credibility of the results. As described above, the NTP monograph was conducted using robust methods that included a thorough evaluation of risk of bias, but like any systematic review of the scientific literature, was reliant on the studies available at the time it was developed. Synthesis of the available evidence as performed by the NTP often requires drawing conclusions in spite of existing uncertainty, such as a lack of ideally conducted studies. The EPA acknowledges that the NTP monograph included studies that they determined to have high risk of bias; the NTP monograph also included studies that had a lower risk of bias. To explore differences in the results among all of these studies, NTP transparently grouped studies by level of bias. In conducting this analysis, which was supported by peer reviewers, NTP found that 18 out of the 19 high-quality (or lower risk of bias) studies reported an inverse association between estimated fluoride exposure and IQ in children, meaning that as fluoride exposure increased, IQ decreased. For purposes of dose-response analysis (as in the EPA's assessment), studies at lower risk of bias are preferred to ensure the most accurate estimates of the relationship between exposure and effect. The EPA's methods for independently evaluating risk of bias are described in Section 6 and include consideration of the concerns raised by commenters.

A.3.3. Dental Fluorosis

Comment 3.3.1: EPA received comments on whether and at what severity dental fluorosis should be considered an adverse health outcome along with comments to predefine and consistently apply criteria for classifying dental fluorosis severity using validated fluorosis indices. The overwhelming majority of commentors stated that mild dental fluorosis is largely a cosmetic concern rather than a structural or functional impairment. A small number of commenters suggested that moderate fluorosis is also a cosmetic effect. Conversely, a few commenters suggested that any severity level of dental fluorosis should be considered adverse.

EPA Response: The EPA agrees that the decisions about the adversity of dental fluorosis and its measurement are important to this assessment and has considered this topic in the protocol. The national and international environmental health community have reached a consensus that severe fluorosis is an adverse health outcome because weakening of tooth enamel can impact tooth function (EFSA Scientific Committee, 2025; Health Canada, 2024; U.S. EPA, 2010a). However, the adversity of other severity levels of dental fluorosis is under debate. In 2006, the National Research Council (NRC) concluded that data were not adequate to determine whether moderate fluorosis was an adverse health effect on the basis of structural or psychological effects (NRC, 2006). However, recent authoritative assessments by EFSA (EFSA Scientific Committee, 2025) and Health Canada (Taher et al., 2024) concluded that moderate fluorosis is adverse and that mild fluorosis is cosmetic. As outlined in Section 2.4, the EPA will consider whether moderate fluorosis is an adverse endpoint for dose-response analysis. The EPA will also predefine and

consistently apply criteria for evaluating evidence from publications on dental fluorosis (Section 6.2.3). The EPA will consider publications that use validated fluorosis indices, including the Dean's Index, Thylstrup-Fejerskov (TF) Index, and the Tooth Surface Index of Fluorosis (TSIF).

Comment 3.3.2: One commenter stated that, although the available evidence supports an association between community water fluoridation and dental fluorosis, it is insufficient to establish causality.

EPA Response: This comment represents a misunderstanding of the process for toxicity assessment development. The hazard determination is not specific to an exposure source (e.g., community water fluoridation) or exposure level associated with that source. The EPA critically reviewed multiple recent health effects documents from national and international health agencies (EFSA Scientific Committee, 2025; Health Canada, 2024; Taher et al., 2024; U.S. EPA, 2010a; NRC, 2006) that reviewed the health effects literature and identified consistency in hazard conclusions for dental fluorosis studies. Steps to determine the dose or exposure levels at which this hazard occurs will be carried out during the dose-response analysis and will be presented in the draft toxicity assessment.

A.3.4. Other Health Effects

Comment 3.4.1: Some commenters were concerned that the EPA would only consider IQ as a neurodevelopmental endpoint, which does not capture potential impacts on the complexity of neurodevelopment and might mischaracterize cognitive and behavioral impact. One of the concerns is that many other neurodevelopmental endpoints might be more sensitive to fluoride. Commenters additionally stated that neurodevelopmental toxicity should be treated as a central endpoint, rather than a secondary or exploratory outcome.

EPA Response: The concerns raised by these commenters are unfounded. The EPA is not limiting the neurodevelopmental endpoints included in the toxicity assessment to IQ. As described in the Refined Populations, Exposures, Comparators, and Outcomes (PECO) criteria (Section 3), the assessment includes <u>any</u> neurodevelopmental outcome (including but not limited to cognition, motor, and behavior). The criteria used to determine the outcomes used for toxicity value derivation are presented in Section 8.1. Neither neurodevelopment nor dental fluorosis are considered "secondary" outcomes, and both will be evaluated fully, based on the best available science, as potential critical effects in the toxicity assessment.

Comment 3.4.2: Some commenters said that the EPA should evaluate cancer, specifically osteosarcoma, as an endpoint. They reported that NRC's committee on fluoride toxicology unanimously concluded that "Fluoride appears to have the potential to initiate or promote cancers," even though the overall evidence is "mixed" (2006). They also pointed to a study that found that boys who drank fluoridated water when they were 6 to 8 years old have a five- to seven-fold greater risk of contracting osteosarcoma by the age of twenty. Lastly, they noted that evidence exists to conclude that fluoride is genotoxic.

EPA Response: As described in the assessment plan (U.S. EPA, 2026), the EPA considered whether cancer should be a focus of this assessment by evaluating cancer conclusions from existing reports developed by national and international health agencies (Section 2.1 of the

assessment plan) and developing a systematic evidence map (SEM) to determine whether additional new data had been published after the existing reports had been released. Existing assessments found that taken together, the epidemiological and animal data did not support an association between fluoride and cancer. The study identified by the commenter that reported the five- to seven-fold greater risk of contracting osteosarcoma for boys is (Bassin et al., 2006), which was previously considered in the Fluoride Human Health Toxicity Assessments published by the (EFSA Scientific Committee, 2025) and (Health Canada, 2010). Although the authors offered a possible mechanistic explanation for why young males were especially susceptible, the authors also cautioned that that their study was an exploratory study and that additional studies are needed to confirm or refute the study findings. The EPA's SEM identified four new epidemiological studies not included in the existing assessments. These studies did not provide new evidence of a positive association between fluoride exposure and cancer. Thus, the overall weight of evidence across published epidemiological studies to date (including Bassin et al. (2006)), which included a variety of geographic regions worldwide and assessed exposures and outcomes across a range of lifestages and study designs, does not support an association between fluoride exposure and carcinogenicity (Section 2.1 and Appendix C of the assessment plan). With regard to toxicological data, including studies that assess genotoxicity as well as cancer initiation and promotion, studies either do not support an association or provide equivocal evidence (Section 2.1 and Appendix C of the assessment plan). Thus, at this time, the EPA determined that reassessment of carcinogenicity with the current dataset would be unlikely to yield a different conclusion from the existing published cancer assessments.

Comment 3.4.3: Some commenters stated that the EPA should consider additional health outcomes beyond those discussed in the assessment plan. Suggested outcomes by at least one commenter included developmental and/or reproductive outcomes (e.g., low birth weight, preterm birth), thyroid endpoints and other endocrine effects (e.g., pineal gland), adult nervous effects, skeletal fluorosis, renal effects, and systemic effects. For some of these effects (e.g., thyroid and developmental), it was suggested that the outcomes often co-occur with or are a possible mechanism for neurodevelopmental outcomes and therefore could strengthen the overall weight of evidence for neurodevelopmental toxicity.

EPA Response: EPA disagrees that additional health outcomes should be included in the toxicity assessment. As discussed in the assessment plan (U.S. EPA, 2026), the EPA considered the results of recent published health effects documents from national and international health agencies to identify the health effects likely to a) have sufficient data to support toxicity value derivation; and b) be a sensitive outcome that will result in a toxicity value that would protect against other health effects in addition to the critical effect selected. Because this approach is used, the Assessment does not require a comprehensive review of all the potential health effects associated with fluoride exposure. A lack of inclusion of additional health outcomes in the assessment does not imply that they are not potential hazards of fluoride exposure. In the case of skeletal fluorosis, which is an accepted hazard, previous reports have indicated that dental fluorosis is likely a more sensitive endpoint to fluoride exposure, and thus a toxicity value based on dental fluorosis is expected to be protective of skeletal fluorosis as well (EFSA Scientific Committee, 2025; U.S. EPA, 2010a; Institute of Medicine, 1997a). Many of the other suggested

outcomes are developing areas of research and additional data are needed to support hazard conclusions and dose-response analysis, as reviewed by EFSA (2025) and Taher et al. (2024).

At this time, review of additional evidence that could support the overall weight of evidence for neurodevelopmental toxicity is not needed because the EPA concluded that neurodevelopmental toxicity is an established hazard based on the critical review of the available information. In other words, review of additional supportive studies for this hazard would not change the assessment conclusions and the agency's decision to conduct a fit-for-purpose toxicity assessment beginning with dose-response analysis for the neurodevelopmental health effects.

For some of the listed outcomes (e.g., renal effects), concerns were raised about whether people with preexisting conditions could be more susceptible to fluoride exposure. Susceptible populations and human population variability in response to fluoride exposure will be considered in the dose-response analysis, as discussed in Section A.4.3 and Comment 5.2.1, and Section 8.

A.4. Assessment Development

A.4.1. Usage of AI Tools

Comment 4.1.1: Some commenters stated that the EPA must clearly document how artificial intelligence (AI) tools are used, emphasizing that "clarifying which tools are to be used, and how, remains of paramount importance." Commenters emphasized that any AI use "needs to be carefully documented so that it can be replicated" including documentation that the EPA "verified the accuracy of all the tools it employed, the prompts it wrote, and what was generated." They asked the EPA to publish "sufficient details to permit independent understanding and review," including the criteria for seed studies, hazard tags, validation steps such as false-negative checks, and reasons for exclusion. Commenters expressed that transparent, reproducible methodology is essential for establishing confidence in the scientific record.

EPA Response: The EPA agrees that transparent, reproducible methodology is essential for developing gold standard scientific products like the Fluoride Human Health Toxicity Assessment. To address the comments, detailed methods on how the EPA implemented AI/automation for literature screening are documented in Section 4 and Appendix C. These sections specify the database search strings, the use of automated/AI tools such as SWIFT-Review filters and SWIFT-Active Screener for title and abstract screening, and the supervised clustering approach with clearly listed seed studies. This ensures transparent and reproducible use of AI tools while upholding scientific integrity. The prompts used for the SWIFT-Review text classification step are provided in Appendix C. This literature screening approach ensures that experts did not receive different results depending on prompt selection or previous queries and allows for a completely reproducible analysis. Additionally, the literature survey results can be transparently reviewed using a user-friendly application that allows the public to search, sort, and follow the studies throughout every step of the review. The EPA will continue to adhere to this transparent approach to documenting methods and results, consistent with the agency's systematic review process (U.S. EPA, 2022b), as the agency develops the Fluoride Human Health Toxicity Assessment.

Comment 4.1.2: Some commenters indicated that AI tools can "struggle to navigate situations without bias," might overlook important studies, and have previously missed "at least five percent of relevant articles" which might lead to exclusion of significant studies. Commenters also noted that they "support the use of such tools when they improve efficiency without compromising comprehensiveness," acknowledging the potential for AI to streamline large-scale literature screening. No matter the position, commenters emphasized that human expertise must remain central so that AI does not "favor studies with similar results" or produce incomplete or biased evidence bases. This includes pairing AI efficiency with safeguards such as random audits, checks for false negatives, and validation against human review to "maximize confidence and reproducibility." Overall, commenters stressed the need to identify AI limitations up front, incorporate strong human oversight into final determinations, and only use AI when supported by rigorous quality control measures that protect scientific integrity.

EPA Response: The EPA agrees that AI tools should be used transparently and with caution. The EPA incorporated safeguards in multiple ways while using automated and AI tools, including a high degree of human oversight. For example, the EPA limited AI bias through the use of two independent human "manual" reviewers when screening titles and abstracts with a senior reviewer to resolve any differences in conclusions. This ensured that AI screening tools were used to assist in sorting and prioritizing the literature, but the EPA human experts ultimately made the final inclusion decisions. Section 4 outlines the steps the EPA took to manage the risk of missed studies. First, the baseline 95% threshold of SWIFT-Active Screener acknowledges a 5% false negative rate (i.e., 5% are incorrectly binned as excluded studies). Therefore, following literature screening, additional checks were conducted using crosswalks with previous assessments, manual human expert review of SWIFT-Review "no-tag" studies, and independent supervised clustering using known true-positive seed studies. Section 4 also outlines the steps of additional levels of review for initially excluded studies to minimize exclusion of relevant studies. The January public comment and future peer review also provide opportunities for commenters to identify studies that might have been inappropriately excluded during screening. Together, these safeguards ensure the final determinations reflect expert judgment and are not skewed by AI algorithms. Each use of AI to support the Fluoride Human Health Toxicity Assessment is transparently documented.

A.4.2. Exposure Measurement

Comment 4.2.1: A commenter noted that it is challenging to determine the precise amount of fluoride that individuals receive from water given that the amount of fluoride fluctuates, intake fluctuates (and is challenging to measure), and actual dose is also based on an individual's height and weight. Relatedly, another commenter noted that intake would vary based on fluoridated water compared with filtered or bottled sources.

EPA Response: The EPA interprets this comment in the context of evaluating health effects studies for dose-response assessment. The EPA agrees that the amount of fluoride that individuals receive from water is challenging to measure and is dependent on intake as well as fluoride concentration, which might vary based on water source. Fluoride concentration in drinking water can be an accurate and direct measure of exposure that can be used as a reliable metric during dose-response analysis. However, it only provides an estimate of a single exposure

source and often requires assumptions about water intake, which can be a source of measurement error and uncertainty. Whether or not individual studies account for water consumption when presenting data will be a consideration during the study evaluation portion of assessment development (see Section 6). Considerations of exposure outside of the context of the evaluation and use of health effects studies to derive an RfD are not within the scope of the fluoride assessment plan and protocol and will be addressed if the agency determines a revision to the fluoride National Primary Drinking Water Rule (NPDWR) is warranted.

Comment 4.2.3: Several commenters noted that exposure to fluoride might come from a variety of sources, including food, beverages, pharmaceuticals, food packaging, dental products, and more, and that water only represents a portion of total exposure. Given this, some of these commenters asked the EPA to clearly specify how it will evaluate total fluoride.

EPA Response: The EPA agrees that fluoride exposure comes from a variety of sources and that this should be considered when assessing fluoride intake. As stated in the assessment plan (U.S. EPA, 2026), measurements of fluoride in urine (Rango et al., 2017) or tissues, such as bone (Rugg-Gunn et al., 2011), teeth (Vieira et al., 2005), or nail clippings (Elekdag-Turk et al., 2019; Rango et al., 2017), are biomarkers of a person's total fluoride intake (Lavalle-Carrasco et al., 2021), accounting for exposure across all sources of fluoride (e.g., drinking water, diet, dental treatments). Criteria for the consideration of exposure metrics during study evaluation are presented in Section 6.2 of the protocol. These biomarkers will be evaluated against other metrics of exposure for their ability to support estimation of oral intake relevant for RfD derivation. Pharmacokinetic modeling approaches will be evaluated for estimation of fluoride intake from biomarkers of exposure to support RfD derivation (see Section 2.4 for additional discussion). Considerations of exposure outside of the context of RfD derivation are not within the scope of the fluoride assessment plan and protocol and will be addressed if the agency determines a revision to the fluoride NPDWR is warranted.

Comment 4.2.4: Many commenters raised concerns with the validity of (and potential exposure misclassification from) using spot blood or urine sampling as proxy for longer-term exposure, given that these biomarkers are subject to intra-individual variability due to influence from recent exposure, hydration status, timing of collection, and/or renal clearance rather than accurately representing typical, average, or cumulative exposure during the relevant critical developmental period for a chronic outcome. The commenters stated that these concerns impact potential conclusions about causality. One commenter asserted that spot urine has not been validated for use in population exposure.

EPA Response: The EPA agrees that use of spot biomarkers to assess fluoride exposure may be subject to potential exposure misclassification due to intra-individual variability or other factors and that this exposure misclassification may impact potential conclusions about causality. However, the EPA disagrees that these concerns invalidate the use of spot urine biomarkers. Overall, exposure to fluoride can be estimated using a variety of methods, such as external measures (e.g., drinking water concentrations) or internal measures (e.g., concentrations in urine, serum, or other biological matrices). Each of these methods has strengths and limitations. For this assessment, the EPA will consider the specific conditions under which internal spot biomarker measures can be potentially valid approaches for characterizing population fluoride

exposure, but strengths and limitations of these exposure metrics will be considered during study evaluation (see Section 6.2 for evaluation criteria). Measurements of fluoride in urine or other tissues are subject to within- and between-person variability due to factors including changing exposure over the lifetime, and inter-individual differences in fluoride elimination. The EPA agrees that fluoride in blood or urine at one point in time might be highly influenced by recent exposures and may not reflect the typical or average exposure of an individual—or exposure at the critical developmental window. Thus, study evaluation will consider whether the reported spot exposure biomarkers reasonably reflect oral intake in an etiologically relevant time period, based on knowledge of fluoride pharmacokinetics and lifestage sensitivities, and how this impacts potential conclusions about causality.

Comment 4.2.5: One commenter noted that 24-hour urine sampling is appropriate for predicting fluoride intake for groups of people but not individuals. Another commenter asserted that while fluoride exposure would correlate with mean spot urine samples, that relationship cannot be assumed for any specific individual mother or their fetus.

EPA Response: The EPA disagrees that 24-hour urine and spot urine sampling are not appropriate for predicting fluoride intake for individuals. A linear relationship between total daily fluoride intake and 24-hour urinary fluoride excretion has been reported in both adults (R^2 = 0.94) and children (R^2 = 0.76) (Villa et al., 2010). Additionally, spot urine samples have been shown to moderately correlate with intake, with a meta-analytic correlation of 0.674 (95% confidence interval [CI]: 0.623–0.725; n = 25 studies) (Eskandari et al., 2023). As such, for the purpose of this assessment, the EPA will consider 24-hour and spot urine samples to be potentially valid approaches of characterizing fluoride human exposure, but strengths and limitations of these and other exposure metrics will be considered during study evaluation. Criteria for the consideration of exposure metrics during study evaluation are presented in Section 6.2.

Comment 4.2.6: A number of commenters noted that the EPA should pre-specify and/or clarify exposure assessment criteria and clearly explain how different methods for (e.g., water concentration, urinary fluoride levels, serum fluoride levels) and timing of exposure assessment across studies will be managed/integrated as well as the approach to handle potential exposure misclassification. Commenters noted that the following clear criteria are crucial to avoid inconsistent comparisons:

- For studies using urinary fluoride (maternal or child), the EPA should pre-specify:
  - Acceptable methods for urine dilution adjustment (creatinine and/or specific gravity) and the implications for bias
  - Whether single spot samples are considered adequate vs preference for repeated measures across pregnancy/early childhood
  - How to handle temporal windows (prenatal vs infancy vs early childhood) without post hoc “critical window” selection
- For studies using water fluoride levels as exposure proxies, the EPA should address:
  - Residential history/mobility

- Water source variability (home vs school, bottled/filtered water)
- Potential misclassification and how it is handled analytically

EPA Response: The EPA agrees that pre-specifying exposure evaluation criteria, including how different methods will be evaluated and how exposure misclassification will be handled, is important to avoid inconsistent comparisons. One of the goals of the protocol is to pre-specify chemical-specific criteria for evaluating studies for use in the toxicity assessment. Exposure to fluoride to support dose-response analysis can be estimated using a variety of methods, such as external measures (e.g., drinking water concentrations) or internal measures (e.g., concentrations in urine, serum, or other biological matrices). For the purpose of this assessment, the EPA will consider both drinking water and internal biomarker measures to be potentially valid approaches of characterizing fluoride human exposure and derive RfDs, but strengths and limitations of the exposure metrics will be considered during study evaluation. The study evaluation criteria used for this assessment are presented in Section 6.2 and these criteria address the issues described by commenters.

A.4.3. Pharmacokinetics

Comment 4.3.1*:* Commenters stressed the need to provide "pharmacokinetic assumptions used to translate internal biomarker concentrations to external intake equivalents." In addition, they emphasized that spot urinary fluoride has "substantial intraindividual variability" and that the "resulting point of departure may not represent a stable estimate of chronic intake." Additionally, one commenter stated that a urine measurement "does not directly quantify time-integrated" dose, which can lead to unstable or misleading benchmark dose estimates when estimating chronic intake.

EPA Response: The EPA agrees that pharmacokinetic assumptions must be clearly stated within the protocol and toxicity assessment in order to address questions related to the use of biomarkers to inform chronic fluoride intake. As part of the robust PK analysis described in Section 6.3, the EPA is evaluating how population- and lifestage-specific models, together with underlying PK data, can be used to characterize interindividual variability in internal biomarkers and improve estimation of the applied dose. Following this analysis, the EPA will determine how to account for intraindividual variability and the impact on a stable estimate of chronic intake. Overall, the models and data used in the assessment will aid in reducing uncertainties surrounding benchmark dose estimates. Criteria for evaluating these considerations in the context of PK model development are provided in Section 6.3.2.

Comment 4.3.2*:* Some commenters stated that internal fluoride dose varies widely across populations because "reduced renal clearance is a well-recognized determinant of increased internal dose," resulting in greater susceptibility of infants, pregnant individuals, those with kidney disease, and others with impaired renal capacity to higher internal fluoride concentrations. They noted that infants "absorb up to 80% of ingested fluoride into their bones" while "healthy adult kidneys excrete only 60% of the fluoride ingested daily." Some commenters also noted that fluoride accumulates in calcifying tissues, with infants and children retaining more fluoride because they excrete less, and that "the fluoride concentration in bone steadily increases over a lifetime." They emphasized that such life-long buildup underscores the

importance of considering retention and systemic distribution when assessing long-term health risks. Overall, commenters argued that risk assessments must account for population-level pharmacokinetic variability through setting appropriate uncertainty factors to account for variability and uncertainty.

EPA Response: The EPA agrees that the internal fluoride dose varies across the population and is the reason that the EPA will conduct a robust analysis of PK data (Section 2.3 of the assessment plan (U.S. EPA, 2026) and Section 6.3 of this protocol) to determine whether a PK model can be developed for fluoride. PK models are the preferred approach to account for variability and sensitivity within a population (U.S. EPA, 2014b), such as lifestage-specific changes in absorption, distribution, metabolism, and excretion (ADME), and are routinely used in the EPA toxicity assessments (U.S. EPA, 2024a, b). Some key features of fluoride ADME are that the fluoride ion is known to preferentially distribute to bone in an age-dependent manner (Richards et al., 1994), does not undergo further biotransformation through metabolism in the body (Institute of Medicine, 1997a), and is mainly excreted in the urine through a lifestage-dependent rate of excretion (Institute of Medicine, 1997a). In conducting this pharmacokinetic analysis (Section 6.3), the EPA will evaluate how population variability and lifestage changes influence fluoride levels relevant to the endpoint of interest, inform reference dose calculations, and determine whether uncertainty factors are needed to further account for differences in susceptibility and long-term exposure within the target population.

Comment 4.3.3*:* Some commenters indicated that water concentration alone is insufficient for exposure assessment because "the same external concentration… can yield materially different systemic exposures across populations and result in different urinary fluoride levels." They argued that when deriving the RfD, intake (mg/kg/day) should serve as the point of departure because water concentration alone "may not reflect aggregate fluoride intake from all relevant sources." Commenters recommended "explicit dose normalization and rigorous PBPK model evaluation" and "intake-based modeling approaches" to ensure an accurate representation of total exposure.

EPA Response: The EPA agrees that fluoride intake (mg/kg/day) should serve as the point of departure and will be used as the basis for deriving the reference dose developed in this assessment. Estimations of fluoride intake will be informed by considering either drinking water fluoride concentrations or measurable biomarkers of fluoride. Section A.4.2 (Exposure Measurement) provides more details on how drinking water measurements will be used. When considering measurable biomarkers, including urine fluoride concentrations, PK models are the preferred method to relate external and internal measures of exposure, which is necessary to derive a reference dose based on internal dose points of departure derived from toxicity, epidemiology, or clinical studies (U.S. EPA, 2006). These biomarkers are representative of total fluoride exposure from all sources and PK models can be used to inform intake from these measurements. As part of pharmacokinetic model evaluation and development outlined in Section 6.3, the EPA will evaluate PK approaches based on the fit to in vivo human data, the accuracy of the biological and mechanistic basis for the model, and the suitability for the model to address the risk assessment context for fluoride (specifically, whether lifestage-specific features of ADME are accurately described). The EPA intends to use the results from the

ongoing pharmacokinetic analysis to estimate intake values from measured water and urinary concentrations while accounting for the different levels of exposure observed across populations and lifestages when deriving the RfD. Consideration of fluoride intake outside of the context of RfD derivation is not within the scope of the Fluoride Human Health Toxicity Assessment plan and protocol and will be addressed if the agency determines a revision to the fluoride NPDWR is warranted.

A.4.4. Methods of the Preliminary Literature Survey

Comment 4.4.1: Some commenters expressed concern that the literature search was conducted in 2024 and requested that "In accordance with Gold Standard Science and its principles on timeliness and comprehensiveness, the EPA's search should be updated to include peer-reviewed publications from 2025 and early 2026 prior to moving forward with the Toxicity Assessment." Additionally, commenters questioned why the assessment plan cited the article by Taylor et al. 2025, stating that "We also request that the EPA provide rationale for inclusion of one paper from 2025 and the exclusion of other 2025 papers in the literature search."

EPA Response: The EPA agrees that the literature search should be updated before the draft toxicity assessment is released and has responded by providing additional details on this in the protocol. The preliminary literature survey was compiled by conducting database searches in PubMed and Web of Science in November 2024, as well as by identifying citations in published reports on fluoride (see protocol Section 4 and Appendix B of the assessment plan (U.S. EPA, 2026)). Updated database searches will be performed during assessment development, including an updated literature search prior to the planned release of the draft toxicity assessment for public comment and external peer review (see protocol Section 4.1.1). The commenters misconstrued the EPA's rationale for citing the publication by the systematic review and meta-analysis by Taylor et al. (2025), which was not identified as an included study in the preliminary literature survey and does not meet the assessment PECO criteria because it does not report original data. Rather, the Taylor et al. (2025) publication was discussed in the assessment plan because it is related to the NTP (2024) report. This was part of the background information on fluoride, which was intended to provide context for the EPA's toxicity assessment. However, some references with early 2025 publication dates were identified in the preliminary literature survey from the database searches because they were made publicly available online toward the end of the 2024 calendar year, and others were identified from citations in the EFSA Scientific Committee (2025) report as part of the crosswalk with existing assessments/reports. Additional new studies that meet the refined PECO criteria will be included in the systematic review process when the literature search is updated.

Comment 4.4.2: Some commenters felt that the assessment plan did not clearly explain the process for study selection and prioritization, stating that "The inclusion of studies spanning multiple decades—without a clear rationale for how they were identified or prioritized—makes it difficult to assess whether the review reflects the most relevant and up-to-date science."

EPA Response: The EPA disagrees that the assessment plan lacked a description of study selection and prioritization rationale. A transparent description of the literature search and screening procedures, and the prioritization of studies for study evaluation and dose-response

analysis, was provided in the assessment plan (U.S. EPA, 2026). Furthermore, a link was provided to an interactive version (the updated version for the protocol is linked here) of the literature flow diagram that allows users to filter specific references to identify the disposition of any individual study. Briefly, the preliminary literature survey was not date limited and all identified studies that met the refined PECO criteria were included. This was intended to avoid omitting potentially relevant studies. The best available science can then be identified using Gold Stand Science principles after reviewing each study that evaluated neurodevelopmental or dental fluorosis after fluoride exposure. For instance, the study of dental fluorosis by Dean et al. (1942) is relevant to the EPA's assessment. While it is one of the oldest publications on fluoride health outcomes, it was used as the basis for recent health effect reference values by EFSA Scientific Committee (2025), Health Canada (2024), and U.S. EPA (2010b). The study by Dean et al. (1942) is also the basis for the current fluoride secondary maximum contaminant level. The EPA plans to update the literature survey before the assessment is released in order to consider the most recent epidemiological data as part of the evidence base.

Comment 4.4.3: One commenter asked for clarification regarding whether non-English studies and gray literature were included, and what the process was for identifying gray literature.

EPA Response: The EPA agrees that the assessment plan (U.S. EPA, 2026) lacked information on the handling of non-English studies in the literature search and has responded by adding this information to the protocol (see Section 4.2.6). The EPA did not specifically conduct a search of gray literature databases for health effect studies, but gray literature was identified as part of the crosswalk with existing fluoride assessments/reports and was screened according to the PECO criteria. This is described in Appendix B of the assessment plan and in Section 4 of the protocol.

A.4.5. Results of the Preliminary Literature Survey

Comment 4.5.1: Some commenters were concerned about the inclusion of potentially biased studies in the preliminary literature inventory, stating that "If the EPA wishes to promote the best available, unbiased, peer-reviewed studies through broad literature searches in scientific databases, it must remove the biased studies included in this review and add numerous, better-conducted studies from the last few years."

EPA Response: EPA disagrees with assertion that studies should be removed from the preliminary literature survey. The preliminary literature survey presented in the assessment plan (U.S. EPA, 2026) included all studies that are relevant to the toxicity assessment because they met the refined PECO criteria. These relevant studies described in the assessment plan have not yet undergone study evaluation for risk of bias and sensitivity. Risk of bias and sensitivity evaluation will be performed on each relevant study independently by the EPA scientists according to the methods described in Section 6.2. The study evaluation results will be used to guide the selection of studies for dose-response analysis, with priority placed on *high* or *medium* confidence studies. Through this systematic review approach, the EPA will ensure that the Fluoride Human Health Toxicity Assessment is based on the best available science and is consistent with Gold Standard Science principles.

Comment 4.5.2: Some commenters stated that the toxicity assessment should focus on studies with ranges of fluoride levels comparable to those in the U.S., stating that "Several studies were

conducted in regions with naturally occurring fluoride concentrations far exceeding the level used in U.S. community water fluoridation (0.7 mg/L), limiting generalizability to U.S. populations. These factors underscore the importance of distinguishing between high-level environmental exposure and optimally fluoridated community water systems." Other potential differences of concern noted between the U.S. and other populations included water consumption patterns, climate, body weight, diet, and co-exposures. One additional commenter stated that dose normalization is necessary when extrapolating neurodevelopmental findings across regions.

EPA Response: The EPA disagrees that only studies with ranges of fluoride comparable to those in the U.S. should be the focus of the toxicity assessment. The goal of the preliminary literature survey in the assessment plan (U.S. EPA, 2026) was to capture all of the available studies that could potentially be used in the toxicity assessment. For this reason, study location and relative exposure level were not considered as part of the inclusion criteria when identifying relevancy of the studies. In addition, studies conducted outside the U.S. are not inherently less valid. Studies with different exposure contexts often provide useful information when developing toxicity assessments for environment contaminants. Studying different exposure scenarios (e.g., high exposure levels, contamination events, sensitive subpopulations, different exposure routes) can be extremely informative in understanding potential adverse effects of these contaminants on human health. During selection of studies for dose-response analysis, relevance to the U.S. population, including relative exposure levels, will be considered, as well as any potential need to conduct dose normalization (Section 8.1).

Comment 4.5.3: One commenter noted that dental fluorosis is featured more prominently compared with neurodevelopmental toxicity in the tables of the assessment plan, which makes it difficult to assess the influence that neurodevelopmental outcomes will have on the overall final conclusions of the EPA's toxicity assessment for fluoride.

EPA Response: This comment has misconstrued the results of the preliminary literature survey that is presented in the assessment plan (U.S. EPA, 2026), which determined that more studies for dental fluorosis were published than neurodevelopmental toxicity. However, this does not mean that dental fluorosis will be weighed more heavily in future conclusions in the EPA's toxicity assessment. Depending on the findings of study evaluation, the EPA will determine a draft RfD for fluoride based on the robustness of the dose-response analyses and the sensitivity of the points of departure for each health effect category (Section 8.1) (U.S. EPA, 2022b) that will be included in the draft toxicity assessment. The future draft assessment will undergo peer review and public comment.

Comment 4.5.4: Regarding the literature tracking tools in Tableau Public, one commenter stated that "We appreciate this transparency and encourage EPA to ensure newly identified studies and commentaries are incorporated into this tracking tool so stakeholders can follow inclusion/exclusion decisions across updates."

EPA Response: Thank you for the positive feedback about the publicly available literature tracking tool. Regarding the request to add newly identified studies with subsequent literature search updates, the EPA agrees that the literature tracking tool should be updated as new studies are identified. The EPA will not add new literature to the existing tracking tool continuously because the link provided in the assessment plan reflects that stage in the process (U.S. EPA,

2026). Rather, at the time of publishing the draft toxicity assessment, the EPA will add a new link to an updated version of the literature tracking tool with the results of the updated literature search. At that time, this will allow the public to search, sort, and follow the complete database of studies throughout every step of the assessment development process.

A.4.6. Additional Sources

Comment 4.6.1: Many commenters suggested additional peer-reviewed publications for the EPA to incorporate in their review.

EPA Response: The EPA has reviewed peer-reviewed references submitted through public comment. References that had not been captured by the literature search were added to the assessment reference workflow (Table 4-1) and were screened according to the assessment PECO (Table 4-3). References that had been captured by the literature search but were excluded during one of the automated prioritization steps (Section 4.2.3) were manually reviewed and included if deemed relevant according to the PECO criteria. Disposition of these studies is available in the interactive literature flow diagram.

Comment 4.6.2: Several commenters included sources of information on fluoride that are not peer reviewed or do not contain primary data (e.g., review articles). These sources include websites, articles, and books.

EPA Response: These sources of information will not be considered in the toxicity assessment because they do not meet the EPA's criteria for peer review and containing primary data. As described in Section 4.1 of the assessment plan (U.S. EPA, 2026) and maintained in Section 4.1 of the protocol, these types of information do not meet the PECO criteria (Table 4-3) and are not considered further for the assessment.

A.4.7. Study or Data Evaluation

Comment 4.7.1: Many commenters requested that the EPA Assessment should be transparent, unbiased, and "grounded in well-established scientific standards," as well as complying with Executive Order 14303. Another comment recommended that the EPA use the Integrated Risk Information System (IRIS) framework for systematic review.

EPA Response: The EPA agrees and is using peer-reviewed systematic review methods in the development of the Fluoride Human Health Toxicity Assessment, which provides a transparent, unbiased assessment based on Gold Standard Science. As discussed in the protocol, the EPA is using published, established methods (U.S. EPA, 2022b) that have been positively peer reviewed by the NASEM. Systematic review enhances consistency, transparency, and reproducibility. It promotes use of the best available science, reduces bias, and upholds scientific integrity.

Comment 4.7.2: Many commenters shared concerns with the quality of studies reporting associations between fluoride exposure and adverse effects, as well as the extent to which these studies can inform the risk of fluoride exposure to the U.S. population. These concerns included potential confounding from socioeconomic factors and education, co-exposure to other neurotoxins (e.g., lead, mercury, arsenic), and nutritional factors, including iodine status and overall nutrition. If these factors are not adequately controlled, commentors suggested the studies be "downgraded for internal validity" and "flagged as limited of U.S. applicability." Concerns

were also raised about evaluation of outcome ascertainment methods, including challenges in measuring neurodevelopmental effects (e.g., IQ) and grading dental fluorosis (e.g., use of validated indices). Commenters also requested the EPA "apply rigorous study quality criteria" and "explain how study quality, data limitations and null or negative findings" will be addressed while one commenter stated that the EPA must reassess studies to ensure they are "of scientific rigor." Another commenter recommended that the EPA "pre-specify an evidence-ranking framework" to identify the most informative studies. One commenter stated that studies "funded or influenced by entities with institutional or policy stakes in fluoridation should undergo transparent and reproducible risk of bias evaluation."

EPA Response: The EPA agrees that the issues raised by the commenters are relevant to the assessment. As described in the protocol, the EPA will address these important concerns through its study evaluation phase of the systematic review, which is the next step in the toxicity assessment process by which the EPA evaluates the strengths and limitations of the relevant studies and considers these factors or concerns. For more details, the study evaluation plan customized for the fluoride assessment is described in Section 6. Section 6.2 describes the criteria and additional considerations applied for each evaluation domain, with specific criteria developed for fluoride exposure and the outcomes included in this assessment. Criteria were developed by a team of subject matter experts. In the case of neurodevelopmental effects, the criteria were developed previously by experts in psychometric testing and have been used in multiple EPA assessments with assessment-specific adaptations (U.S. EPA, 2024c, 2020, 2019c). All studies considered for dose-response analysis will undergo study evaluation, and as described in Section 8.1, studies of "high" or "medium" confidence will be prioritized for further analysis.

Comment 4.7.3: One commenter recommended that the EPA specify evaluation and exclusion criteria for dose-response suitability. They recommended that the EPA report the rationale for exclusion from dose-response consideration. Another commenter suggested that the criteria to determine suitability for dose-response analysis must include "exposure metric validity, reproducibility across analytic specifications, and internal consistency across related publications derived from the same cohort."

EPA Response: The EPA agrees with these concerns and notes that these concerns are addressed through the EPA's systematic review approach to developing toxicity assessments (U.S. EPA, 2022b). Section 8.1 describes the considerations used to prioritize and select studies for dose-response analysis, including factors such as exposure metric validity and analytic methods. Rationales for excluding studies from dose-response analysis will be transparently documented in the assessment.

Comment 4.7.4: Some commentors recommended the EPA mostly rely on cohort studies, rather than cross-sectional or observational data, to understand the longer-term effects of fluoride exposure.

EPA Response: The EPA disagrees with this comment's suggestion to exclude studies based on study design. The EPA's systematic review approach to developing human health toxicity assessments does not prioritize studies based on type of study alone, but rather the totality of how the study was conducted and its informativeness to the research question. As described in Section 5, one of the considerations for prioritizing studies for study evaluation is use of a

measure of fluoride exposure that is collected prospectively and which represents exposure during an etiologically relevant period, or a fluoride measure that is collected cross-sectionally but can reasonably be considered representative of exposure during an etiologically relevant period. This consideration generally results in prioritization of cohort studies but allows for suitable studies with other study designs to be identified.

Comment 4.7.5: Several commenters mentioned the expertise required for this assessment. One commenter recommended the inclusion of expertise in "psychometrics and cognitive test methodology," "developmental psychology or pediatric neuropsychology," "dental and oral health science," and "probabilistic aggregate exposure assessment." Other commenters stated that "peer reviewers must have objective expertise on fluoride and neurocognition" and "should involve trusted and well-established researchers."

EPA Response: The EPA agrees with the overarching recommendation that the assessment be developed by technical experts. The EPA's large, multidisciplinary team includes a wide range of expertise relevant to and necessary for developing a fit-for-purpose toxicity assessment based on Gold Standard Science. In addition, the team relies on well-established, peer-reviewed, published guidance and methods (e.g., U.S. EPA (2022b, 2012)) and other relevant documents developed by subject-specific experts to inform decisions (e.g., NIEHS (2022)). The EPA will engage additional expertise as needed, including during peer review.

Comment 4.7.6: One commenter stated that the "EPA should consider identifying and grouping identical or substantially identical submissions, publicly reporting the number of duplicate comments received, clearly distinguishing mass generated template comments from individualized submissions, and implementing review mechanisms that ensure substantive scientific comments are not obscured."

EPA Response: The EPA's process is consistent with the commenter's recommendation. The EPA used a structured approach to categorize the public comments received. Each comment was reviewed thoroughly by multiple EPA scientists to ensure that all substantive and unique comments were considered and not overlooked. Mass comment campaigns with substantially identical information were identified and catalogued. The number of similar comments in each campaign identified is available in the docket for this action (EPA-HQ-OW-2025-3823 at www.regulations.gov).

A.4.8. RfD Derivation – Benchmark Response and Uncertainty Factors

Comment 4.8.1: Commenters identified several groups of people who are potentially more susceptible to the effects of fluoride exposure, including pregnant women and fetuses, formula-fed infants and children, older adults, people with higher water consumption, and individuals with specific health conditions (e.g., reduced kidney function, immune or endocrine dysfunction, fluoride hypersensitivity, arthritis, inflammatory bowel disease). Commenters asked the EPA to explicitly consider these groups when selecting uncertainty factors, and where appropriate, to conduct population-specific risk assessments. One commenter stated that the EPA's decision in 2010 not to evaluate fluoride intake in infants less than 6 months of age resulted in the "most exposed and most vulnerable group" being excluded from the determination of safe exposure

levels. Commenters also identified several factors that might modify individual susceptibility to the effects of fluoride exposure, including genetic variation, nutritional status, co-exposures to other chemicals, and total fluoride intake from all sources. One commenter highlighted the disproportionate impact that a reduction in IQ could have on certain individuals, stating that a 5-point reduction might have more severe consequences for children at the lower end of the distribution.

EPA Response: The EPA agrees that some populations are potentially more susceptible to fluoride exposure and will consider information about these populations when conducting the toxicity assessment for fluoride. Specifically, the EPA will identify sensitive lifestages and populations from reviewing the relevant literature and then use this information when conducting dose-response assessment (Section 8) and selecting uncertainty factors (Section 8.2.2). Uncertainty factors will be considered and selected when developing candidate RfDs, in line with the EPA's relevant guidance documents (e.g., U.S. EPA (2002b)). This step will include consideration of variation in susceptibility among the members of the human population (i.e., inter-individual or intraspecies variability) from reviewing the available, relevant studies. The response to Comment 4.3.2 provides additional discussion on the role of pharmacokinetic factors as they relate to potentially susceptible populations. Potential differential impact of harmful effects will be considered during selection of the benchmark response (BMR).

Comment 4.8.2: Some commenters urged the EPA to "follow its published procedures by applying the default intraspecies uncertainty factor of 10 absent data sufficient to justify departure" and to ensure that the "RfD is based solely on risk-related principles, and not policy concerns." They stated that, in their view, policy considerations influenced the selection of uncertainty factors in previous assessments and urged the EPA to clearly identify all uncertainty factors and provide transparent justification for their selection. Another commenter stated that additional uncertainty factors might be required to address database uncertainty (e.g., where specific subpopulations show greater sensitivity or where studies necessary to assess sensitive populations are absent or inadequate).

EPA Response: This comment is premature for the assessment's current stage. The EPA will follow its published, final human health risk assessment guidance and best practices document on selecting the intraspecies uncertainty factor (U.S. EPA, 2002b). Rationales for the selected uncertainty factor will be presented in the draft assessment, and the public will have an opportunity to comment on UF selections at that time.

Comment 4.8.3: One commenter asked that the EPA clearly document the effect of exposure misclassification on benchmark response estimates.

EPA Response: This comment is premature for the assessment's current stage. The BMR is based on biological and statistical consideration. However, the Benchmark Dose Technical Guidance (U.S. EPA, 2012) does not mention exposure measurement error. Rationales for selected benchmark responses will be presented in the draft assessment, and the public will have an opportunity to comment on the benchmark response selections at that time. The EPA will follow final agency human health risk assessment guidance in selecting the benchmark response for RfD derivation (U.S. EPA, 2012).

Comment 4.8.4: One commenter stated that if the EPA uses moderate dental fluorosis as the adverse health outcome, a benchmark response (BMR) of 10% is more appropriate than the BMR of 0.5% for severe dental fluorosis.

EPA Response: This comment is out of scope for review of the assessment plan (U.S. EPA, 2026). When the EPA publicly releases the draft toxicity assessment for fluoride, the public will have an opportunity to comment on BMR selections. As the EPA conducts the dose-response analysis, the EPA will follow agency guidance in selecting the benchmark response for RfD derivation (U.S. EPA, 2012).

Comment 4.8.5: A commenter recommended that benchmark dose modeling be prioritized, where data permit. Another requested that the EPA "carefully evaluate dose-response relationships."

EPA Response: The EPA agrees with the comment. Consistent with the EPA guidance, the EPA prioritizes benchmark dose modeling where data permits in order to consider the full dose-response relationship in estimating points of departure (refer to the *Benchmark Dose Technical Guidance* (U.S. EPA, 2012)). The EPA will present the details of the benchmark dose modeling methods and results in the draft toxicity assessment, a document that will undergo peer review and public comment.

Comment 4.8.6: Another commenter encouraged the EPA to apply benchmark dose modeling to studies that reported biomarkers of exposure. They recommended that, in those cases, the EPA document 1) "The pharmacokinetic assumptions used to translate internal biomarker concentrations to external intake equivalents," 2) "The potential impact of exposure misclassification on benchmark response estimates," and 3) "The stability of BMD and BMDL estimates across alternative covariate specifications, interaction terms, and treatment of influential observations." They also recommended transparent presentation of "benchmark modeling inputs, alternative model fits, and sensitivity analyses" in the protocol and draft assessment.

EPA Response: The EPA agrees and the planned approach for deriving the draft RfD is consistent with the recommendations. Specifically, the EPA will evaluate the suitability of PK models for extrapolation of biomarkers of exposure (such as blood and/or urine fluoride concentrations) to total average daily exposure (external intake equivalents) (Section A.4.3). Exposure estimates for dose-response analysis might be based on either external or internal measures of exposure (Section 8.1). All analysis and modeling decisions and assumptions will be documented transparently in the draft toxicity assessment.

Comment 4.8.7: Some commenters recommended that the analysis include stratification of studies based on fluoride exposure levels and, in some comments, to pre-specify these strata in the systematic review protocol. The recommendation was for these strata to be used to exclude or de-emphasize studies with higher exposure from dose-response analysis, as well as other analyses, such as subgroup and sensitivity analysis. A commenter stated that "dose relevance is critical for risk interpretations and for informing standards under the Safe Drinking Water Act."

EPA Response: The EPA agrees with the idea of considering exposure levels during dose-response analysis but disagrees with the idea of pre-specifying strata of exposures. As part of the

systematic review approach being used to develop the toxicity assessment, PECO-relevant studies with informative dose-response data will undergo prioritization based on key study attributes, such as study confidence (based on the results of study evaluation), exposure paradigm, subject selection, confounding, measurements of exposure and health outcome, study design, and size. All study attributes are examined from the perspective of the effect on dose-response modeling. Studies with exposures near the range of typical environmental human exposures will be preferred and studies with a broad exposure range and multiple exposure levels are preferred to the extent that they can provide information about the shape of the exposure-response relationship (see the EPA *Benchmark Dose Technical Guidance*, §2.1.1) and facilitate extrapolation to more relevant (generally lower) exposures. (Table 7-2 in (U.S. EPA, 2022b)).

Comment 4.8.8: Commenter suggested that the EPA seek access to the underlying data particularly from key studies to "ensure that analysis was appropriate" and "reproduce the findings reported."

EPA Response: The EPA disagrees with the comment. The EPA's dose-response assessment methods can be performed on the modeling results provided in publications or by the study authors and do not require raw underlying data details. This approach has successfully been employed in previous the EPA assessments (e.g., PFOA and PFOS toxicity assessments (U.S. EPA, 2024a, b)). Furthermore, raw underlying epidemiology data from individuals typically contains information that is confidential. Data use agreements used by researchers and study participants are legally binding contracts that generally restrict sharing of data with outside entities to protect participant confidentiality and personally identifiable information.

A.4.9. Peer Review and Engagement with the Public and Other Agencies.

Comment 4.9.1: Several commenters requested that the protocol be released for public comment.

EPA Response: Because the EPA incorporated public feedback on the assessment plan in the protocol, the protocol is not being released for public comment. Protocol Sections 2–4 contain revisions to the assessment plan based on considering public comment. The EPA will provide another opportunity for public comment on the draft toxicity assessment. At that time, the public can comment on both the methods described in the protocol and how the EPA executed the protocol in the draft assessment conclusions. The EPA's toxicity assessment is being developed consistent with published peer-reviewed systematic review methods (U.S. EPA, 2022b).

Comment 4.9.2: Many commenters stated that the EPA should enlist the National Academies of Sciences, Engineering, and Medicine (NASEM) to peer review the toxicity assessment and/or protocol prior to finalization. One of these commenters stated that peer review by NASEM would ensure compliance with federal standards outlined in the Office of Management and Budget (OMB) Peer Review Bulletin of 2004. Additionally, one commenter stated that the EPA should consider commissioning an independent review by the NASEM to evaluate the human health evidence on fluoride in parallel with the ongoing Fluoride Human Health Toxicity Assessment by the EPA.

EPA Response: These comments address the next steps in the EPA's process leading up to a new Fluoride Human Health Toxicity Assessment. Following the assessment plan (U.S. EPA,

2026), the EPA developed a protocol that the agency will follow in developing the draft toxicity assessment. The EPA disagrees that both the protocol and toxicity assessment should be peer reviewed by the NASEM prior to finalization. Rather, the EPA plans to have the toxicity assessment undergo peer review according to the agency's peer review policy as described in the *EPA Peer Review Handbook (2015)*, which will include review by the Science Advisory Board (SAB). Of note, the SAB contains components similar to aspects of the NASEM emphasized by the OMB Peer Review Bulletin of 2004, including the review of balance, conflict of interest, and independence of members who conduct the peer review. Additionally, the toxicity assessment is anticipated to contain frameworks consistent with previous recommendations by the NASEM on the established systematic review methods (U.S. EPA, 2022b) that the EPA will use for the toxicity assessment, as explained in the protocol.

Comment 4.9.3: One commenter requested that the "EPA and FDA [U.S. Food and Drug Administration] should ensure their toxicological and safety evaluations of fluoride are consistent across regulatory frameworks." Another commenter expressed sentiment for the EPA to send the rulemaking task over to the Center for Drug Evaluation and Research (CDER) of FDA. Another commenter specifically requested that the EPA coordinate with "CDC [Center for Disease Control and Prevention], HRSA [Health Resources and Services Administration], IHS [Indian Health Services], and state oral health programs" to "align water policy with public health outcomes." Furthermore, one commenter requested that the EPA coordinate with the National Institute of Dental and Craniofacial Research (NIDCR) to "explicitly integrate NIDCR-coordinated data sharing and longitudinal-cohort linages into the assessment protocol."

EPA Response: These comments offer recommendations for regulatory actions or interagency coordination related to fluoride which is beyond the scope of the Assessment Plan document (U.S. EPA, 2026). As explained in the assessment plan, the EPA is coordinating closely with HHS about the human health toxicity assessment development. The results of this toxicity assessment will also be used to inform a whole-of-government approach to Making America Healthy Again, including Centers for Disease Control and Prevention (CDC) recommendations regarding adding fluoride to drinking water.

Comment 4.9.4: Several commenters requested that "the EPA include its Office of Applied Science and Environmental Solutions (OASES) and their Children's Health Protection Division in this Assessment Plan and Toxicity Assessment to ensure proper research protocol and ample expertise in children's health." Similarly, multiple commenters were concerned about infrastructure for conducting objective and thorough scientific reviews due to the EPA's recent reorganization, particularly reorganization of the Office of Research and Development (ORD). These commenters were "also concerned by the lack of consultation with other federal oral health and research professionals on this review."

EPA Response: The EPA agrees with the comment about coordination within the agency. The EPA's toxicity assessment for fluoride is being developed by experts in human health toxicity assessments from across the EPA.

A.5. Other Topics

A.5.1. Benefits of Fluoride Exposure

Comment 5.1.1: Commenters disagreed with the EPA's decision to exclude beneficial health effects from the assessment plan and urged the EPA to evaluate the full body of evidence, including studies demonstrating the safety and health benefits of community water fluoridation. One commenter noted that prior assessments conducted by the EPA have considered benefits when deriving uncertainty factors and reference doses and stated that it would be inappropriate to derive a reference dose using standard "conservative assumptions" without considering the beneficial health effects. Additional commenters recommended that the final toxicity assessment explicitly acknowledge the public health benefits of community water fluoridation, even if they are not within the scope of the toxicity assessment. In contrast, one commenter supported the decision to exclude health benefits during the toxicity-assessment phase, stating that it is "critical that risk assessment decisions be made independent of these considerations."

EPA Response: The EPA disagrees that beneficial health effects of fluoride exposure or specifically benefits of community water fluoridation should be included in the toxicity assessment. As the Safe Drinking Water Act (SDWA) prohibits EPA from requiring the addition of any substance to drinking water for any preventive health care purpose, the potential health benefits of fluoride, such as cavity prevention, are not within the scope of the agency's statutory authorities and this toxicity assessment. As stated in the assessment plan (U.S. EPA, 2026), the overall objective of the toxicity assessment is to develop an RfD for oral fluoride exposure. An RfD is defined as an estimate of daily oral exposure to humans that is likely to be without an appreciable risk of deleterious effects during a lifetime. Therefore, the EPA is prioritizing review of evidence related to adverse health effects in this assessment. The EPA will follow its established human health risk assessment guidance and best practices (U.S. EPA, 2022b, 2002b), including for uncertainty factor application, to develop the toxicity assessment. This toxicity assessment will be used to inform future decisions about potential revisions to the existing federal regulations on fluoride in drinking water under the SDWA, thus addressing concerns about children's health including developmental impacts. The results of this toxicity assessment will also be used to inform a whole-of-government approach to Making America Healthy Again, including Centers for Disease Control and Prevention (CDC) recommendations regarding adding fluoride to drinking water. The decision on whether or not to add fluoride to drinking water is made on a state or local basis; EPA does not make recommendations on adding fluoride to drinking water.

Comment 5.1.2*:* Several commenters urged the EPA to clarify how public health benefits will be integrated in subsequent risk management and regulatory decisions. Commenters warned that without an explicit framework for incorporating benefits at later stages, the toxicity assessment could present an incomplete and misleading picture of the public health implications. Accordingly, commenters recommended that the EPA evaluate potential health and economic impacts, including effects on vulnerable populations, prior to finalizing any regulatory changes. One commenter urged EPA to ensure that the toxicity assessment outputs are compatible with benefit–risk modeling, including dose-response relationships for adverse effects, exposure

distributions for sensitive populations, and uncertainty bounds suitable for comparison with caries-prevention thresholds. Some commenters also shared their views pertaining to the relative risks and benefits of water fluoridation policies and the potential future impact that the human health toxicity assessment or potential future rulemaking could have on those policies.

EPA Response: EPA addresses the topic of inclusion of beneficial effects in the toxicity assessment as the response to Comment 5.1.1. Comments on benefits inclusion, health risk reduction and cost analysis in potential future rulemaking and the relative risk and benefits of water fluoridation policies are out of scope for the assessment plan (U.S. EPA, 2026) and protocol currently under consideration. The responses to comment 4.3.2 and 5.2.1 relate to how vulnerable populations are planned to be addressed in the toxicity assessment. There will be additional opportunities to comment on the draft toxicity assessment regarding how toxicity assessment outputs are presented.

A.5.2. Potential Future Rulemaking

Many commenters expressed their views on potential future regulatory actions by the EPA under the Safe Drinking Water Act (SDWA) to address fluoride in drinking water. *While the EPA has made no decision to revise the fluoride National Primary Drinking Water Regulation (NPDWR)*, and while these comments are outside the scope of the Fluoride Human Health Toxicity Assessment and its Preliminary Assessment Plan and Literature Survey (U.S. EPA, 2026), EPA has responded to comments below for clarification purposes.

Comment 5.2.1*:* Commenters recommended that the EPA establish a specific value for the potential MCLG or recommended the use of specific safety factors for the potential MCLG. Commenters referenced the EPA's prior SYR determinations for fluoride, including past decisions not to revise the standard, and emphasized the need for similar consideration in the current assessment, or urged the EPA to not accelerate the SYR timeline. Commenters raised concerns about how exposure in susceptible populations, including formula-fed infants, would be considered in developing a potential new MCLG for fluoride and suggested the EPA consider additional data on sensitive subpopulations in future Six-Year Reviews.

EPA Response: Comments pertaining to development of an MCLG are not responsive to the assessment plan or protocol. The final toxicity assessment will serve as an updated scientific foundation that can inform future steps under the SDWA. If EPA determines at a future date to revise the fluoride NPDWR and propose and promulgate a revised MCLG, there will be an opportunity to comment on the MCLG value and the process for MCLG development. The EPA disagrees that it is inappropriate to perform a toxicity assessment for fluoride at this time. Section 1412(b)(9) of the SDWA requires the EPA to, not less often than every six years, review existing NPDWRs and, if appropriate, revise them. This process is called Six-Year Review (SYR) and was last completed in July of 2024 (SYR4, U.S. EPA, 2024f). The EPA notes that prior SYR determinations for fluoride were conducted in accordance with the SDWA and reflected the best available science at the time of those reviews. The EPA also notes that the fact that an NPDWR was not selected for revision in prior SYR determinations means only that EPA believes that regulatory changes to a particular NPDWR are not appropriate at the time of that action, and future reviews may identify information that leads to an initiation of the revision

process. While the EPA determined that revisions to the fluoride NPDWR were not possible at the time of the fourth SYR, they also stated that "following publication of the final NTP report, EPA will consider the systematic review and meta-analysis conclusions regarding developmental neurotoxicity to inform the agency's future development of a health effects assessment for fluoride" (SYR4, U.S. EPA, 2024f). While the next comprehensive analysis of new scientific information on potential health risks of fluoride in drinking water would be required by 2030 (six years after the fourth SYR), the EPA aims to deliver updated science in the form of this toxicity assessment to the public sooner, while maintaining rigorous review and quality controls. Consideration of exposure in susceptible populations as it pertains to the toxicity assessment is addressed in Comment 5.2.1. Comments pertaining to future six-year reviews are not responsive to the assessment plan or protocol and are thus out of scope.

Comment 5.2.2: A commenter raised the question about the potential for other contents of tap water to be harmful.

EPA Response: This comment is not responsive to the assessment plan or protocol and is thus out of scope.

Comment 5.2.3: One commenter felt that action on fluoride in drinking water would be better addressed by the Toxic Substances Control Act (TSCA), rather than the SDWA.

EPA Response: This comment does not pertain to the assessment plan or protocol and is therefore considered out of scope.

Comment 5.2.4*:* One commenter stated that the "EPA regulates drinking water but disclaims preventive health use." Commenters expressed concerns about the EPA's and the FDA's regulatory frameworks in regard to fluoride in drinking water. Some commenters recommended that the EPA align its risk communication and guidance for drinking water fluoride exposures with FDA determinations that caution or restrictions are needed for ingestible fluoride products at comparable doses and for specific age groups.

EPA Response: Comments regarding the addition of fluoride to drinking water and FDA's regulatory framework for fluoride are outside the scope of this toxicity assessment and the EPA's purview. The EPA does not make recommendations on adding fluoride to drinking water because the SDWA prohibits the EPA from requiring the addition of any substance to drinking water for preventive health care purposes (Section 1412(b)(11)). Therefore, by law, the EPA cannot prescribe fluoride for preventive health use. The SDWA grants the EPA the authority to establish regulations that limit the occurrence of contaminants in drinking water. Comments regarding the broad alignment of interagency risk communication between the EPA and FDA are outside of the scope of this assessment action. Comments regarding the coherence between the toxicity assessment and FDA determinations are premature, as the conclusions of toxicity assessment have not been made.

A.5.3. Assessment Development Process

Comment 5.3.1: Two commenters requested that the EPA extend the public comment period.

EPA Response: The EPA considered the request for an extension of the public comment period for the assessment plan (U.S. EPA, 2026). The EPA posted a copy of the assessment plan on January 28, 2026, and accepted comments through February 27, 2006. The EPA considers the

30-day comment period to have been an appropriate and meaningful opportunity to provide transparency and gather early feedback and elected not to extend the public comment period. Additionally, the EPA did not provide an extension of the public comment period because there will be additional opportunities for public comment at later stages of the toxicity assessment development process, which will be announced at a future date.

Comment 5.3.2: One comment recommended haste in assessment development, specifically by reducing repetitive reviews and bureaucracy.

EPA Response: The ongoing toxicity assessment is a fast-track effort to expedite EPA's next fluoride health assessment under the Safe Drinking Water Act schedule (see Comment 6.2.1), while adhering to Gold Standard Scientific methods and radical transparency. The assessment will reflect a comprehensive, current, and accurate review of the best available scientific data on fluoride exposure and health effects, conducted with external peer review and public input to inform potential revisions to federal drinking water regulations that bring the nation closer to evidence-based standards that protect future generations. EPA will use Gold Standard Scientific practices, draw on expertise across the federal government, and will not prejudge any outcomes of the assessment.

Comment 5.3.3: Some commenters stated that the assessment plan was unclear about the current status of the assessment and that "some sections imply the assessment is already completed."

EPA Response: EPA agrees that there was a lack of clarity regarding the status of the assessment. As part of scoping and problem formulation, the literature search and screening stages of the assessment were performed prior to publication of the assessment plan and the protocol (see Section 4 of the assessment plan (U.S. EPA, 2026) and Section 4 of this protocol). Conducting these steps in advance was a critical step toward publishing the assessment plan and developing the complete protocol document. The remaining steps of the systematic review and dose-response assessment which have yet to be completed are outlined in Sections 5–8.

Comment 5.3.4: A commenter recommended revising the assessment plan prior to starting the protocol.

EPA Response: The EPA disagrees that it is necessary to revise the assessment plan prior to development of the protocol. The EPA incorporated feedback received on the assessment plan (U.S. EPA, 2026) in the development of the protocol, which includes the majority of the content from the assessment plan (Sections 2, 3, and 4).

A.5.4. Risk Communication

Comment 5.4.1: One commenter emphasized the need for "clear communication distinguishing hazard identification from risk assessment."

EPA Response: EPA agrees with this comment. The protocol includes additional information on the risk assessment process, including describing the key steps of hazard identification, dose-response analysis, exposure assessment, and risk characterization. Section 1 describes how this toxicity assessment fits into the larger risk assessment process. Additionally, the EPA published fact sheets describing the steps in the systematic review and the purpose of the assessment plan, protocol, and toxicity assessment.

Comment 5.4.2: Commenters urged the EPA to clearly communicate scientific uncertainty and limitations so that stakeholders can understand what the evidence shows as well as where and why the evidence is limited or inconsistent.

EPA Response: The EPA agrees and will include a thorough discussion of scientific uncertainty in the draft toxicity Assessment. This will include both qualitative discussion of scientific decisions and their limitations as well as the application of quantitative adjustment for uncertainty (i.e., uncertainty factors), such as accounting for human variability and identified database deficiencies. The assessment plan (U.S. EPA, 2026) and the protocol are the first steps in ensuring transparency and providing documentation of how the evidence will be identified, evaluated, and assessed, including the topics mentioned by commenters (Sections 4, 6, and 8). The draft assessment will continue with this effort by documenting any updates to the protocol methods and transparently describing all analyses and decisions.

Comment 5.4.3: Several commenters requested that the EPA "clearly state that any derived reference dose for fluoride is simply a threshold to avoid adverse effects and does not negate the well-documented caries-prevention benefits and safety profile at recommended levels."

EPA Response: The EPA will provide clear explanations of any derived reference dose and how it should be interpreted in the draft assessment. As stated in the assessment plan (U.S. EPA, 2026) and protocol, descriptions of the health benefits of fluoride exposure are outside the scope of the toxicity assessment which focuses on adverse health effects after fluoride exposure. However, the EPA responded to comments regarding health benefits for clarification purposes in Section A.5.1.

Comment 5.4.4: Some commenters stated that the assessment plan risks conflating "hazardous levels of fluoride" (i.e., above the maximum contaminant level and secondary standards of 4.0 mg/L and 2.0 mg/L) with the U.S.-recommended community water fluoridation levels (0.7 mg/L), potentially mischaracterizing the current fluoridation levels as dangerous. Commenters recommended that the EPA use clear, evidence-based communication that distinguish between findings relevant to community water fluoridation levels and those at substantially higher exposure levels.

EPA Response: As described in the response to comment 3.1.2, the EPA's acceptance of hazard for adverse health effects does not identify or discuss a particular level of fluoride. Specifically, this section does not discuss the current community water fluoridation level as hazardous. Comments about the community water fluoridation level for health benefits are beyond the scope of the toxicity assessment. The EPA has not yet characterized the dose or exposure levels at which the hazards occur. As the EPA is following a systematic review approach and agency guidance for human health risk assessment (e.g., U.S. EPA (2022b, 2012)), this characterization will be evidence-based and transparently documented in the draft assessment.

Comment 5.4.6: Some commenters emphasized that the EPA's fluoride assessment has the potential to influence state drinking water standards. They urged the EPA to provide clear, plain-language guidance to help state and local decision-makers interpret and communicate the results of the assessment since any ambiguity could lead to the misinterpretation or misapplication of the toxicity values and enforcement standards. In the same vein, one commenter said that "states,

municipalities, water utilities, and the public routinely rely on EPA drinking water standards as assurances of safety." The commenter recommended that the EPA explicitly state the scope and limits of the fluoride assessment and that "MCLGs [maximum contaminant level goals] and related benchmarks do not constitute affirmative safety determinations."

EPA Response: The EPA appreciates the potential impact of the final toxicity assessment on a potential future revision of the fluoride drinking water regulation. In the draft toxicity assessment, the EPA will include clear language regarding interpretation of draft toxicity values, as well as in other plain-language materials that will be published online.

Comment 5.4.7: Commenters recommended that the EPA "avoid the use of normative or persuasive language—such as describing certain sources as 'authoritative'."

EPA Response: The draft assessment will adhere to standard scientific language found in other assessment documents published by the agency. In the assessment plan (U.S. EPA, 2026), the word "authoritative" was used to denote national and international health agencies or organizations that routinely develop toxicity assessments. "Authoritative" was not intended to be normative or persuasive but instead reflect standard, objective, and well-accepted classifications of the information source. However, this language has been revised throughout the protocol.

Comment 5.4.8: One commenter highlighted that there was no federally mandated fluoride monitoring in private wells and that the assessment plan should acknowledge this regulatory gap as well as whether these populations are appropriately represented in the literature. Another commenter noted that there are no regular measurements being made of the levels of fluoride in urine, blood, bones, hair, or nails in the general population or among sensitive groups in the population.

EPA Response: The toxicity assessment is being developed to provide an updated scientific foundation that can inform future steps under the Safe Drinking Water Act (SDWA). Comments about policies regarding fluoride monitoring in private wells and/or as regular monitoring across the country are outside the scope of this assessment. When final, the toxicity assessment and any potential future revisions to the NPDWR may provide useful information to communities and homeowners with well water.

A.5.5. Recent Fluoride Litigation Under the Toxic Substances Control Act

Comment 5.5.1: Some comments raised concerns that the EPA's approach described in the assessment plan is not consistent with the position represented by the agency in a recent court proceeding (Food & Water Watch, Inc., et al. v. United States Environmental Protection Agency).

EPA Response: Comments addressing this litigation, which involves a challenge to EPA's denial of a citizen group petition, submitted pursuant to Section 21 of the Toxic Substances Control Act (TSCA), requesting that the agency ban the addition of fluoride to public drinking water, and assertions regarding statements made by the agency therein, are outside the scope of this assessment action. These comments addressing assertions EPA made regarding health effects of fluoride misapprehend the action that the agency put out for comment here. Namely,

the assessment plan is not a toxicity assessment; it does not contain conclusions regarding harmful human health effects of fluoride or determine the levels of fluoride exposure associated with harmful health effects. Such conclusions about fluoride human health effects (e.g., toxicity values) will be released as part of the forthcoming draft toxicity assessment.

Comment 5.5.2: One commenter stated that the EPA should "align with the 2024 court ruling (Food & Water Watch, Inc., et al. v. United States Environmental Protection Agency) deeming 0.7 mg/L 'unreasonable risk,'" based on NTP's findings. Similarly, other commenters advocated for the EPA to use NTP's findings to recommend immediate rule development.

EPA Response: The EPA disagrees with the comment that EPA should rely on a decision from ongoing litigation and/or the NTP monograph findings to immediately develop a rule. These comments represent a misunderstanding of the EPA process for developing a rule under the SDWA, as well as the posture of this action that the agency released for comment. The EPA's assessment plan clearly describes the purpose and findings from critically reviewing the NTP monograph (U.S. EPA, 2026). The NTP monograph and other recent, final documents on fluoride health effects developed by national and international health agencies provide the basis for establishing the neurodevelopmental hazard for fluoride. To develop the new EPA Fluoride Human Health Toxicity Assessment, the EPA will follow agency guidance and rely on the best available science to determine the level below which adverse effects are not anticipated. Following external peer review and public comment, this assessment will be the foundational scientific evaluation needed to inform the agency's future steps for review of the fluoride regulation under the Safe Drinking Water Act.

Appendix B. Quality Assurance

Quality assurance (QA) activities performed by the U.S. Environmental Protection Agency (EPA) ensure that environmental data are of sufficient quantity and quality to support the agency's intended use. The work within this report was conducted under the agency's QA program for environmental information.

Agency-wide, the EPA Quality System provides the framework for planning, implementing, documenting, and assessing work performed by the agency, and for carrying out required quality assurance and quality control (QA/QC) activities. Additionally, the Quality System covers the implementation of the EPA Information Quality Guidelines (U.S. EPA, 2002a). This protocol follows all agency guidelines to ensure a high-quality document.

Within the EPA, quality management plans (QMPs) and quality assurance project plans (QAPPs) are developed to ensure that all agency materials meet a high standard for quality. All staff generating and reviewing information for this protocol followed detailed procedures documented in the Programmatic Quality Assurance Project Plan for Collection, Use, and Analysis (including Model Application) of Existing Data (P-QAPP), under EPA Contract 68HERC23D0003, approved April 2023 and Quality Management Plan For Human Health Microbiology and Toxicology Regulatory Support (QMP), under EPA Contract 68HERC23D0003, approved May 2023. The procedures outlined in the QMP and P-QAPP are in compliance with EPA's Quality Program policies, standards, and guidance (https://www.epa.gov/quality/quality-program-directives) and other related policies as cited in the P-QAPP and QMP. In addition, ICF maintains literature search and screening standard operating procedures that are compliant with the Office of Water's QMP. The protocol underwent review by EPA staff and management.

Appendix C. Literature Search and Screening

C.1. Database Search Strings for Fluoride

Sources used to identify chemical synonyms include the following: CompTox (only "Valid" and "Good" synonyms), ChemIDPlus, Cameo Chemicals, medical subject headings (MeSH) Dictionary (PubMed only), previous EPA/OW literature searches (SYR4, Contaminant Candidate List 5, Contaminant Candidate List 6), and consultation with subject matter experts. The chemical identifiers used were limited to the Chemical Abstracts Service Registry Number (CASRN), the EPA CompTox Distributed Structure-Searchable Toxicity Substance Identifier (DTXSID), and the Unique Ingredient Identifier (UNII) used by the United States Food and Drug Administration.

Table C-1. Database Search Strategies for Fluoride

Database	Search Terms	Number of References[a,b]
Web of Science	TS=("16984-48-8" OR "DTXSID9049617" OR "florinate" OR "florinated" OR "fluoridate" OR "fluoridated" OR "Fluoride" OR "Fluorides" OR "fluorinate" OR "fluorinated" OR "Fluorine" OR "fluorosis" OR "fluridate" OR "fluridated" OR "FLUORIDE-ION" OR "Hydrofluoric acid, ion(1-)" OR "hydrofluosilicic acid" OR "Perfluoride" OR "Sodium silicofluoride" OR "UNII-Q80VPU408O")	**Total references:** 235,996 **Filtered for English language:** 227,384
PubMed	("16984-48-8"[rn] OR "16984-48-8"[tiab] OR "DTXSID9049617"[tiab] OR "florinate"[tiab] OR "florinated"[tiab] OR "fluoridate"[tiab] OR "fluoridated"[tiab] OR "Fluoride"[tiab] OR "Fluorides"[mh] OR "Fluorides"[tiab] OR "fluorinate"[tiab] OR "fluorinated"[tiab] OR "Fluorine"[tiab] OR "fluorosis"[tiab] OR "Fluorosis, Dental"[mh] OR "fluridate"[tiab] OR "fluridated"[tiab] OR "FLURORIDE-ION"[tiab] OR "Hydrofluoric acid, ion(1-)"[tiab] OR "hydrofluosilicic acid"[tiab] OR "Perfluoride"[tiab] OR "Sodium silicofluoride"[tiab] OR "UNII-Q80VPU408O"[tiab])	**Total references:** 102,955 **Filtered for English language:** 94,516
Total number of unique references identified (English language) in initial database searches: 268,646		

[mh] = MeSH term; [rn] = Registry number; [tiab] = title-abstract.

[a]Initial database searches were conducted on November 1, 2024 with no date limits.

[b]"Total references" represents all references identified using the search terms. "Filtered for English language" represents the number of references remaining after filtering the results to prioritize English language studies.

C.2. Automated Prioritization Approaches for Health Effects Studies

Prior to manual review of studies by subject matter experts, a variety of natural language processing and machine-learning techniques can be applied to large evidence bases to efficiently identify studies most likely to be relevant. The automated tools are well-established for the systematic review of health effects literature and have been used to develop several final EPA toxicity assessments (e.g., U.S. EPA (2025e, 2024e, 2023c, 2019c)).

C.2.1. Title-Abstract Prioritization

Prior to manual title-abstract screening, the 268,093 unique studies identified in the literature search underwent prioritization using machine learning-based artificial intelligence (AI) and automated approaches.

C.2.1.1. SWIFT-Review Filtering

SWIFT-Review filters were used to automatically tag studies by evidence stream. Of these, 118,629 references received prioritized evidence stream tags (Section 4.2.3.1), 39,623 studies received non-prioritized evidence stream tags, and 109,845 studies received no tag. Additional prioritization tools (i.e., topic extraction and supervised clustering) were then applied to studies in the prioritized evidence stream and "no tag" groups to identify studies most likely to be informative to the toxicity assessment. The 39,623 studies that received non-prioritized evidence stream tags were excluded from further consideration.

C.2.1.2. Health Effects Tagged Studies: Supervised Clustering and Topic Extraction

For the 118,629 references that received health effects tags, supervised clustering and topic extraction approaches were applied. For supervised clustering, 42 populations, exposures, comparators, and outcomes (PECO)-relevant studies were selected from previous EPA assessments of fluoride and from manual review of the list of studies that received evidence stream tags during SWIFT-Review filtering to serve as seed studies (Table C-2).

Table C-2. Seed Studies Used for Supervised Clustering

Reference	Title
(Erickson et al., 1976)	Water fluoridation and congenital malformations: No association
(Erickson, 1980)	Down syndrome, water fluoridation, and maternal age
(Maurer et al., 1990)	Two-year carcinogenicity study of sodium fluoride in rats
(Das et al., 1994)	Toxic effects of chronic fluoride ingestion on the upper gastrointestinal tract
(Jacqmin et al., 1994)	Components of drinking water and risk of cognitive impairment in the elderly
(Skotowski et al., 1995)	Risk factors for dental fluorosis in pediatric dental patients
(Heindel et al., 1996)	Developmental toxicity evaluation of sodium fluoride administered to rats and rabbits in drinking water
(Pendrys et al., 1996)	Risk factors for enamel fluorosis in a nonfluoridated population
(Mascarenhas and Burt, 1998)	Fluorosis risk from early exposure to fluoride toothpaste

Reference	Title
(Kumar and Swango, 1999)	Fluoride exposure and dental fluorosis in Newburgh and Kingston, New York: policy implications
(Li et al., 2001)	Effect of long-term exposure to fluoride in drinking water on risks of bone fractures
(Acharya, 2005)	Dental caries, its surface susceptibility and dental fluorosis in South India
(Bataineh and Nusier, 2006)	Impact of 12-week ingestion of sodium fluoride on aggression, sexual behavior, and fertility in adult male rats
(Bera et al., 2007)	Neurofunctional effects of developmental sodium fluoride exposure in rats
(Reddy et al., 2007)	Suppression of male reproduction in rats after exposure to sodium fluoride during early stages of development
(Rocha-Amador et al., 2007)	Decreased intelligence in children and exposure to fluoride and arsenic in drinking water
(Kivrak and Kars, 2012)	Effects of fluoride on anxiety and depression in mice
(Picco et al., 2014)	The effect of chronic treatment with fluoride on salivary activity, tooth, and bone in spontaneously hypertensive rats
(Yeung, 2014)	Some beneficial effect on root caries from use of higher concentration fluoride toothpaste (5000 ppm F)
(Charugundla et al., 2015)	Comparative effect of fluoride, essential oil and chlorhexidine mouth rinses on dental plaque and gingivitis in patients with and without dental caries: a randomized controlled trial
(Li et al., 2015)	Pathological changes and effect on the learning and memory ability in rats exposed to fluoride and aluminum
(Niu et al., 2016)	Changes in liver antioxidant status of offspring mice induced by maternal fluoride exposure during gestation and lactation
(Saha et al., 2016)	Musculoskeletal problems and fluoride exposure: A cross-sectional study among metal smelting workers
(Barberio et al., 2017a)	Fluoride exposure and indicators of thyroid functioning in the Canadian population: implications for community water fluoridation
(Bashash et al., 2017)	Prenatal fluoride exposure and cognitive outcomes in children at 4 and 6-12 years of age in Mexico
(Barberio et al., 2017b)	Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation
(Lu et al., 2017)	Sodium fluoride causes oxidative stress and apoptosis in the mouse liver
(Valdez Jiménez et al., 2017)	In utero exposure to fluoride and cognitive development delay in infants
(Bartos et al., 2018)	Alterations in the memory of rat offspring exposed to low levels of fluoride during gestation and lactation: Involvement of the α7 nicotinic receptor and oxidative stress
(Khandare et al., 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
(Sun et al., 2018)	Effects of fluoride on SOD and CAT in testis and epididymis of mice
(Bartos et al., 2019)	Effects of perinatal fluoride exposure on short- and long-term memory, brain antioxidant status, and glutamate metabolism of young rat pups
(Green et al., 2019)	Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada
(Liu et al., 2019)	Fluoride exposure and pubertal development in children living in Mexico City
(Riddell et al., 2019)	Association of water fluoride and urinary fluoride concentrations with attention deficit hyperactivity disorder in Canadian youth
(Soto-Barreras et al., 2019)	Effect of fluoride in drinking water on dental caries and IQ in children
(Russ et al., 2020)	Aluminium and fluoride in drinking water in relation to later dementia risk
(Sharashenidze et al., 2020)	Dental fluorosis prevalence, severity and associated risk factors in pre-school-aged children residing in fluoride deficient regions of Georgia

Reference	Title
(Till et al., 2020)	Fluoride exposure from infant formula and child IQ in a Canadian birth cohort
(Sowanou et al., 2022)	Association between osteoarthritis and water fluoride among Tongyu residents, China, 2019: A case-control of population-based study
(Kampouri et al., 2023)	Associations of gestational and early-life exposure to toxic metals and fluoride with a diagnosis of food allergy or atopic eczema at 1 year of age
(Lee et al., 2024)	Association of fluoride exposure with disease burden and neurodevelopment outcomes in children in South Korea

The seed studies represent a range of study types (e.g., epidemiology studies, animal toxicology studies, and studies that included mechanistic findings) and exposure routes (e.g., drinking water and oral exposures in addition to topical application to the oral cavity). The supervised clustering results are provided below in Table C-3.

Table C-3. Supervised Clustering Results for Fluoride Literature Following SWIFT-Review Filtering

Ensemble Score	Counts	Seeds in Cluster
0	88,889	2
1	14,861	4
2	6,161	7
3	1,862	5
4	1,961	2
5	4,184	15
6	711	7
Grand Total	**118,629**	**42**

After reviewing the results, the 14,879 studies with ensemble scores 2–6 were forwarded for title-abstract screening, as these scores included seed studies and were predicted to be relevant by at least two models.

Ensemble scores 0 and 1 (i.e., those generally assumed to be least relevant) also contained seed studies and were further prioritized using topic extraction in Litstream™, which clusters references into groups based on similarity of keywords used in the titles and abstracts. Topic extraction identified 25 groups of common terms across the studies in the supervised clustering ensemble scores 0 and 1 (Table C-4).

The EPA identified topic clusters 2, 3, 9, 11, 15, 16, 19, 21, 23, and 25 as groups containing potentially relevant studies based on the topic extraction terms. Importantly, the six seed studies contained in the supervised clustering ensemble scores 0 and 1 were captured in topic clusters 2 and 21. The EPA also identified topic clusters 1 and 4 as containing potentially relevant mechanistic terms.

Therefore, the EPA advanced the 48,772 studies with health effects tags that were identified in bins 1, 2, 3, 4, 9, 11, 15, 16, 19, 21, 23, and 25 to title and abstract screening. In total, 63,651 studies that received hazard tags in SWIFT-Review were prioritized for title-abstract screening.

Table C-4. Topic Extraction of Fluoride Supervised Clustering Ensemble Scores 0 and 1 Using 25 Bins

Topic Cluster Number	Number of References	Topic Keywords
1	*6,330*	*['activity', 'enzyme', 'protein', 'inhibited', 'protease', 'purified', 'cells', 'inhibitor', 'ph', 'serine', 'inhibitors', 'inhibition', 'proteins', 'acid', 'mm', 'dependent', 'phosphatase', 'binding', 'presence', 'rat']*
2	**6,505**	**['children', 'dental', 'caries', 'health', 'oral', 'oral health', 'dental caries', 'year', 'prevalence', 'years', 'fluorosis', 'preventive', 'care', 'teeth', 'risk', 'school', 'study', 'age', 'dental fluorosis', 'prevention']**
3	**1,185**	**['contrast', 'contrast enhanced', 'ultrasound', 'ceus', 'enhanced', 'enhanced ultrasound', 'contrast enhanced ultrasound', 'patients', 'sonovue', 'liver', 'contrast agent', 'enhancement', 'ultrasonography', 'lesions', 'imaging', 'perfusion', 'diagnosis', 'sonography', 'ultrasound ceus', 'agent']**
4	*1,766*	*['cyclase', 'adenylate', 'adenylate cyclase', 'cyclase activity', 'adenylate cyclase activity', 'activity', 'stimulated', 'adenyl', 'adenyl cyclase', 'rat', 'stimulation', 'membranes', 'basal', 'enzyme', 'activation', 'hormone', 'isoproterenol', 'adrenergic', 'receptor', 'gtp']*
5	1,294	['glass ionomer', 'ionomer', 'glass', 'cement', 'resin', 'ionomer cement', 'glass ionomer cement', 'cements', 'release', 'materials', 'gic', 'restorative', 'ionomer cements', 'glass ionomer cements', 'resin modified', 'modified glass', 'resin modified glass', 'restorative materials', 'fuji', 'modified']
6	3,707	['18', 'pet', 'imaging', 'positron', 'tomography', 'positron emission', 'emission tomography', 'positron emission tomography', 'fluorine 18', 'emission', 'labeled', '18f', 'brain', 'uptake', 'vivo', 'fluorine', 'pet imaging', '18 labeled', 'radiochemical', 'emission tomography pet']
7	1,188	['fuel', 'membranes', 'proton', 'membrane', 'fuel cell', 'fuel cells', 'exchange', 'nafion', 'conductivity', 'proton exchange', 'proton conductivity', 'cell', 'sulfonated', 'exchange membrane', 'polymer', 'poly', 'membrane fuel', 'proton exchange membrane', 'cm', 'exchange membranes']
8	10,035	['structure', 'stable', 'crystal', 'temperature', 'calculations', 'fluorine', 'complexes', 'phase', 'energy', 'hydrogen', 'structures', 'atoms', 'state', 'degrees', 'formation', 'properties', 'ray', 'reaction', 'molecular', 'using']
9	**2,335**	**['enamel', 'dental enamel', 'human', 'dental', 'human enamel', 'teeth', 'surface', 'vitro', 'tooth', 'uptake', 'caries', 'effect', 'fluorosis', 'remineralization', 'acid', 'enamel surface', 'calcium', 'demineralization', 'tooth enamel', 'surface enamel']**
10	8,735	['surface', 'films', 'properties', 'film', 'oxide', 'water', 'fluorine', 'coating', 'using', 'high', 'fluorinated', 'doped', 'spectroscopy', 'degrees', 'ray', 'electron', 'coatings', 'prepared', 'based', 'layer']
11	**1,179**	**['caries', 'dental caries', 'dental', 'prevention', 'prevention dental', 'prevention dental caries', 'caries prevention', 'children', 'fluorine', 'prophylaxis', 'dental caries prevention', 'dental caries children', 'caries children', 'fluorides', 'water', 'prevention caries', 'caries fluorine', 'fluoridation', 'prophylaxis dental', 'dental caries fluorine']**
12	2,737	['pvdf', 'piezoelectric', 'vinylidene', 'poly', 'poly vinylidene', 'flexible', 'sensor', 'energy', 'polyvinylidene', 'based', 'output', 'polymer', 'trfe', 'sensors', 'mechanical', 'wearable', 'film', 'high', 'ferroelectric', 'vdf']
13	3,351	['nm', 'emission', 'luminescence', 'doped', 'red', 'fluorescence', 'green', 'upconversion', 'ions', 'light', 'excitation', 'er3', 'emitting', 'color', 'detection', 'blue', 'laser', 'fluorescent', 'optical', 'based']
14	2,353	['solar', 'solar cells', 'cells', 'perovskite', 'efficiency', 'pscs', 'pce', 'perovskite solar', 'power conversion', 'perovskite solar cells', 'conversion', 'conversion efficiency', 'photovoltaic', 'power conversion efficiency', 'device', 'power', 'layer', 'polymer', 'performance', 'based']

Topic Cluster Number	Number of References	Topic Keywords
15	**20,762**	**['effect', 'sodium', 'bone', 'fluorine', 'human', 'effects', 'study', 'treatment', 'patients', 'cells', 'clinical', 'use', 'using', 'vitro', 'used', 'acid', 'results', 'caries', 'studies', 'rat']**
16	**716**	**['retinal', 'eyes', 'detachment', 'macular', 'vitrectomy', 'retinal detachment', 'gas', 'surgery', 'visual', 'macular hole', 'tamponade', 'acuity', 'visual acuity', 'hole', 'patients', 'postoperative', 'sf6', 'pars plana', 'plana', 'pars']**
17	1,601	['sensitized', 'sensitized solar', 'dye', 'dye sensitized', 'dye sensitized solar', 'solar', 'sensitized solar cells', 'solar cells', 'tio2', 'dsscs', 'fto', 'cells', 'efficiency', 'oxide', 'counter', 'tin', 'tin oxide', 'doped', 'fluorine doped', 'doped tin']
18	9,346	['compounds', 'synthesis', 'fluorinated', 'derivatives', 'activity', 'reaction', 'fluorine', 'nmr', 'synthesized', 'new', 'substituted', 'compound', 'alpha', 'fluoro', 'acid', 'novel', 'group', 'binding', 'analogues', 'series']
19	**1,290**	**['pet', 'ct', 'pet ct', 'bone', '18', 'naf', 'naf pet', 'patients', 'prostate', '18 naf', 'cancer', 'imaging', '18f', 'naf pet ct', 'prostate cancer', 'metastases', 'uptake', 'tomography', '18 naf pet', 'bone metastases']**
20	1,865	['membrane', 'membranes', 'pvdf', 'flux', 'water', 'separation', 'fouling', 'polyvinylidene', 'pvdf membrane', 'surface', 'distillation', 'membrane distillation', 'pore', 'performance', 'rejection', 'polyvinylidene pvdf', 'poly', 'contact', 'pvdf membranes', 'oil']
21	**6,012**	**['enamel', 'group', 'groups', 'control', 'study', 'dentifrice', 'significant', '05', 'treatment', 'dentin', 'test', 'effect', 'containing', 'significantly', 'teeth', 'remineralization', 'plaque', 'lesions', 'specimens', 'caries']**
22	3,856	['fdg', 'pet', 'fdg pet', 'ct', 'tomography', '18', 'pet ct', 'patients', 'positron emission', 'positron', 'emission tomography', 'positron emission tomography', 'fluorine 18', 'fdg pet ct', '18 fdg', 'fluorodeoxyglucose', 'emission', '18 fluorodeoxyglucose', 'uptake', 'imaging']
23	**860**	**['fluorides', 'topical', 'topical application', 'caries', 'application', 'topical fluorides', 'prevention', 'use fluorides', 'effect', 'use', 'caries prevention', 'dental', 'fluorides caries', 'prevention fluorides', 'effect fluorides', 'effect topical', 'topical applications', 'caries prevention fluorides', 'fluorides dental', 'application fluorides']**
24	3,727	['lithium', 'batteries', 'electrolyte', 'li', 'ion', 'capacity', 'high', 'electrolytes', 'lithium ion', 'ion batteries', 'electrochemical', 'cathode', 'battery', 'performance', 'metal', 'cycling', 'anode', 'solid', 'cycles', 'stability']
25	**1,017**	**['osteoporosis', 'bone', 'treatment osteoporosis', 'treatment', 'therapy', 'postmenopausal', 'fractures', 'calcium', 'fracture', 'sodium', 'therapy osteoporosis', 'vitamin', 'patients', 'women', 'postmenopausal osteoporosis', 'vertebral', 'bone mineral', 'bone mass', 'bone loss', 'mineral']**

Bold font indicates topic clusters that appear to include studies meeting populations, exposures, comparators, and outcomes (PECO) criteria based on the listed terms (N = 40,676 studies). *Italicized font* indicates topic clusters that appear to include mechanistic information based on the listed terms (N = 8,096 studies).

C.2.1.3. "No Tag" Studies: Iterative Topic Extraction

For the 109,845 "no tag" studies, an iterative topic extraction approach was applied. The prioritization was carried out in three phases:

1) All "no tag" references were placed into 25 groups ("topics") according to the keywords commonly used in the titles and abstracts (Table C-5). These 25 topics were reviewed to identify relevant terminology. Topic Cluster 12 included terms such as "caries," "fluorosis," and "teeth" that were presumed to refer to studies with relevant information.

Additionally, Clusters 3, 12, and 14 also included studies identified in the crosswalk with published assessments, suggesting that these clusters included some relevant literature. These three clusters were further refined as described below:

1) To further refine the literature to forward to title-abstract screening, an additional targeted topic extraction was conducted of the 21,815 studies from clusters 3, 12, and 14 (Table C-6). Clusters 1, 6, 9, 11, 13, 15, 18, 20, and 25 (n = 12,666 studies) from this additional targeted extraction appeared to include relevant information based on the listed terms. The subset of studies from Clusters 1, 6, 9, 11, 13, 15, 18, and 20 that did not overlap with studies identified in the crosswalks (n = 3,216) were forwarded to SWIFT-Active Screener for title-abstract screening.
2) After the topic extraction in Table C-5, many studies (n = 9,450) were included in Cluster 25, which appeared to have relevant terms. An additional topic extraction was conducted on Cluster 25 to identify potentially relevant studies (Table C-7). Studies in Clusters 2, 6, 7, 11, 13, 16, 17, 19, 22, 24, and 25 that did not overlap with studies identified in the crosswalks (n = 5,353) were forwarded to SWIFT-Active Screener for title-abstract screening.

The other topic clusters in Table C-2 did not include relevant terminology and were excluded from further consideration. Therefore, following the iterative topic clustering processes, the EPA advanced a total of 8,569 “no tag” studies to title-abstract screening.

Table C-5. Topic Extraction Results for Fluoride Literature with No Tag in SWIFT-Review

Topic Cluster	Number of References	Keywords/Topic Signature
1	4,036	'catalyzed', 'reaction', 'coupling', 'aryl', 'palladium', 'cross coupling', 'synthesis', 'mild', 'alkenes', 'radical', 'cross', 'functional', 'palladium catalyzed', 'conditions', 'copper', 'yields', 'bond', 'reactions', 'group', 'developed'
2	5,771	'complexes', 'crystal', 'structure', 'ii', 'ray', 'complex', 'ligands', 'ligand', 'structures', 'crystal structure', 'iii', 'angstrom', 'metal', 'single', 'compounds', 'single crystal', 'diffraction', 'group', 'nmr', 'coordination'
3	**15,123**	**'sodium', 'effect', 'calcium', 'water', 'ion', 'acid', 'lithium', 'study', 'solutions', 'effects', 'reactions', 'new', 'reaction', 'synthesis', 'ions', 'activity', 'chloride', 'potassium', 'formation', 'fluorine'**
4	4,187	'pvdf', 'piezoelectric', 'polyvinylidene', 'composites', 'composite', 'polymer', 'properties', 'poly', 'poly vinylidene', 'polyvinylidene pvdf', 'vinylidene', 'films', 'dielectric', 'phase', 'mechanical', 'film', 'electrical', 'based', 'nanocomposites', 'high'
5	2,264	'poly', 'vinylidene', 'poly vinylidene', 'pvdf', 'blends', 'polymer', 'crystallization', 'phase', 'methacrylate', 'poly methyl', 'methyl methacrylate', 'electrolyte', 'poly methyl methacrylate', 'hfp', 'ionic', 'conductivity', 'blend', 'electrolytes', 'temperature', 'pmma'
6	5,375	'fluorinated', 'synthesis', 'synthesis fluorinated', 'compounds', 'new', 'properties', 'derivatives', 'polymers', 'new fluorinated', 'highly fluorinated', 'novel', 'partially fluorinated', 'partially', 'reactions', 'highly', 'fluorinated compounds', 'novel fluorinated', 'analogs', 'liquid', 'acids'
7	1,445	'ferroelectric', 'trifluoroethylene', 'trfe', 'vinylidene trifluoroethylene', 'vdf', 'vdf trfe', 'vinylidene', 'copolymer', 'poly vinylidene trifluoroethylene', 'polarization', 'films', 'poly', 'poly vinylidene', 'copolymers', 'field', 'polymer', 'trifluoroethylene vdf', 'trifluoroethylene vdf trfe', 'dielectric', 'switching'

Topic Cluster	Number of References	Keywords/Topic Signature
8	547	'dot center', 'dot center dot', 'center dot center', 'center dot', 'dot', 'center', 'hydrogen', 'interactions', 'center dot hydrogen', 'dot hydrogen', 'hydrogen bonds', 'bonds', 'crystal', 'pi', 'dot hydrogen bonds', 'bond', 'center dot pi', 'dot pi', 'intermolecular', 'complexes'
9	6,812	'films', 'surface', 'film', 'plasma', 'oxide', 'fluorine', 'deposition', 'etching', 'spectroscopy', 'layer', 'ray', 'deposited', 'doped', 'tin', 'silicon', 'using', 'electron', 'si', 'fluorine doped', 'tin oxide'
10	1,986	'glasses', 'glass', 'doped', 'ions', 'optical', 'er3', 'emission', 'glass ceramics', 'properties', 'absorption', 'ceramics', 'spectra', 'heavy metal', 'transition', 'luminescence', 'zrf4', 'baf2', 'crystallization', 'heavy', 'fluorescence'
11	1,691	'18', '18f', 'fluorine 18', 'pet', 'imaging', 'tomography', 'positron', 'positron emission', 'emission tomography', 'positron emission tomography', 'fluorine', 'labeled', 'emission', '18 labeled', 'radiochemical', 'fdg', 'labeling', 'pet ct', 'ct', 'synthesis'
12	**1,466**	**'caries', 'dental', 'enamel', 'fluorosis', 'dental caries', 'dental fluorosis', 'prevention', 'effect', 'uptake', 'topical', 'caries prevention', 'fluoridated', 'dental enamel', 'lesions', 'plaque', 'remineralization', 'prevention dental', 'prophylaxis', 'prevention dental caries', 'teeth'**
13	9,265	'synthesis', 'reaction', 'alpha', 'beta', 'yields', 'derivatives', 'trifluoromethyl', 'substituted', 'compounds', 'reactions', 'acid', 'corresponding', 'fluorinated', 'fluoro', 'amino', 'products', 'ketones', 'group', 'conditions', 'good'
14	**5,226**	**'fluorine', 'fluorine containing', 'containing', 'compounds', 'fluorine compounds', 'fluorine chemistry', 'chemistry', 'synthesis', 'reactions', 'synthesis fluorine', 'organic', 'atoms', 'reaction', 'fluorine atoms', 'synthesis fluorine containing', 'organic fluorine', 'chlorine', 'properties', 'aromatic', 'substituted'**
15	453	'ionomer', 'glass ionomer', 'release', 'glass', 'cements', 'ionomer cements', 'glass ionomer cements', 'cement', 'release glass', 'restorative', 'ionomer cement', 'glass ionomer cement', 'release glass ionomer', 'resin', 'glass ionomers', 'ionomers', 'restorative materials', 'materials', 'resin modified', 'modified glass'
16	2,362	'membrane', 'membranes', 'pvdf', 'water', 'flux', 'separation', 'surface', 'fouling', 'polyvinylidene', 'performance', 'pvdf membrane', 'poly', 'pore', 'pvdf membranes', 'prepared', 'properties', 'polyvinylidene pvdf', 'contact', 'hollow', 'poly vinylidene'
17	2,759	'fluorides', 'earth', 'metal fluorides', 'earth fluorides', 'alkaline earth', 'alkaline', 'alkaline earth fluorides', 'metal', 'alkali', 'synthesis', 'rare', 'rare earth', 'transition metal fluorides', 'reactions', 'new', 'structure', 'acid fluorides', 'transition', 'transition metal', 'uranium'
18	5,132	'laser', 'doped', 'crystals', 'optical', 'nm', 'luminescence', 'emission', 'fiber', 'ions', 'excitation', 'absorption', 'single', 'wavelength', 'crystal', 'energy', 'power', 'spectra', 'earth', 'radiation', 'high'
19	1,840	'hydrogen', 'anhydrous hydrogen', 'anhydrous', 'hydrogen bonding', 'bonding', 'fluorine', 'hydrogen bond', 'liquid hydrogen', 'fluorine hydrogen', 'bond', 'reaction', 'liquid', 'hydrogen fluorine', 'hydrogen bonds', 'hf', 'reactions', 'bonds', 'water', 'acid', 'complexes'
20	2,264	'nmr', '19', 'resonance', 'magnetic resonance', 'magnetic', 'nuclear', 'nuclear magnetic', 'nuclear magnetic resonance', 'fluorine', '19 nmr', 'spin', '19f', 'spectra', 'spectroscopy', 'nmr spectra', 'chemical', 'fluorine 19', 'shifts', '19 nuclear', 'nmr spectroscopy'
21	716	'spectrum', 'rotational', 'nu', 'vibrational', 'microwave', 'constants', 'spectra', 'microwave spectrum', 'rotation', 'state', 'band', 'cm', 'states', 'ground', 'infrared', 'resolution', 'transitions', 'nu nu', 'rotational spectrum', 'structure'
22	1,904	'anion', 'anions', 'fluorescence', 'fluorescent', 'binding', 'receptor', 'detection', 'sensing', 'based', 'colorimetric', 'recognition', 'selective', 'receptors', 'sensor', 'ions', 'chemosensor', 'synthesized', 'ion', 'nmr', 'selectivity'
23	20,436	'surface', 'high', 'temperature', 'properties', 'using', 'based', 'results', 'materials', 'fluorinated', 'degrees', 'fluorine', 'phase', 'used', 'low', 'performance', 'different', 'adsorption', 'study', 'effect', 'process'
24	1,689	'determination', 'determination fluorine', 'electrode', 'selective electrode', 'ion', 'selective', 'ion selective', 'fluorine', 'method', 'ion selective electrode', 'potentiometric', 'spectrophotometric',

Topic Cluster	Number of References	Keywords/Topic Signature
		'method determination', 'using', 'spectrophotometric determination', 'samples', 'detection', 'chromatography', 'analysis', 'spectrometry'
25	5,096	'calculations', 'energy', 'initio', 'ab initio', 'ab', 'energies', 'experimental', 'molecular', 'calculated', 'fluorine', 'bond', 'state', 'results', 'vibrational', 'electronic', 'atoms', 'molecules', 'theoretical', 'theory', 'states'

Bold font indicates topic clusters that appear to include studies that meet the populations, exposures, comparators, and outcomes (PECO) criteria based on the listed terms or included studies identified in the crosswalk with published assessments (N = 21,815 studies).

Table C-6. Targeted Topic Extraction Results for Fluoride Literature with No Tag in SWIFT-Review in Clusters 3, 12, and 14

Topic Cluster	Number of References	Keywords/Topic Signature
1	**550**	**'water', 'fluorosis', 'drinking', 'drinking water', 'fluoridation', 'water fluoridation', 'levels', 'skeletal fluorosis', 'skeletal', 'endemic', 'content', 'endemic fluorosis', 'content drinking', 'content drinking water', 'removal', 'water supply', 'levels drinking', 'supply', 'levels drinking water', 'problem'**
2	601	'ions', 'acid', 'solutions', 'fluorine ions', 'acid solutions', 'aqueous', 'effect', 'phosphoric', 'presence', 'phosphoric acid', 'fluorine', 'extraction', 'nitric', 'solution', 'nitric acid', 'presence ions', 'hydrofluoric acid', 'hydrofluoric', 'sulfuric', 'production'
3	355	'catalytic', 'modified', 'catalysts', 'activity', 'catalytic activity', 'catalyst', 'fluorine', 'reactions', 'reaction', 'using', 'active', 'catalytic properties', 'properties', 'site', 'acid', 'effect', 'lewis', 'supported', 'alumina', 'presence'
4	446	'fluoro', 'phosphorus', 'fluorine', 'synthesis', 'compounds', 'properties', 'phosphorus fluorine', 'chemistry', 'containing', 'new', 'fluorine phosphorus', 'atom', 'reaction', 'reactions', 'derivatives', 'substituted', 'activity', 'chemical', 'substituents', 'deoxy'
5	2,109	'fluorine', 'carbon', 'chlorine', 'atomic', 'atomic fluorine', 'fluorine substituted', 'substituted', 'reaction', 'carbon fluorine', 'atom', 'silicon', 'fluorine atom', 'reactions', 'study', 'synthesis', 'bond', 'fluorine chlorine', 'fluorine bond', 'effect fluorine', 'oxygen'
6	**392**	**'caries', 'dental caries', 'dental', 'prevention', 'caries prevention', 'prevention dental', 'prophylaxis', 'prevention dental caries', 'fluorides', 'effect', 'caries prophylaxis', 'preventive', 'use', 'fluorosis', 'artificial', 'artificial caries', 'root', 'caries preventive', 'dental caries prevention', 'root caries'**
7	465	'atoms', 'fluorine atoms', 'fluorine', 'substitution', 'fluorine substitution', 'reactions', 'reactions fluorine atoms', 'reactions fluorine', 'reaction', 'reaction fluorine atoms', 'substitution fluorine', 'reaction fluorine', 'aromatic', 'nucleophilic', 'effects fluorine', 'effects', 'oxygen', 'compounds', 'effects fluorine substitution', 'atom'
8	1,232	'fluorine containing', 'containing', 'fluorine', 'synthesis', 'synthesis fluorine containing', 'synthesis fluorine', 'compounds', 'new', 'aromatic', 'new fluorine containing', 'new fluorine', 'polymers', 'derivatives', 'beta', 'properties', 'reaction', 'reactions', 'acids', 'containing polymers', 'fluorine containing polymers'
9	**154**	**'osteoporosis', 'fluorophosphate', 'treatment', 'treatment osteoporosis', 'di isopropyl', 'therapy', 'isopropyl', 'di isopropyl fluorophosphate', 'isopropyl fluorophosphate', 'therapy osteoporosis', 'di', 'sodium', 'bone', 'glaucoma', 'osteoporosis sodium', 'diisopropyl fluorophosphate', 'diisopropyl', 'postmenopausal', 'treatment glaucoma', 'postmenopausal osteoporosis'**
10	348	'fluorine compounds', 'compounds', 'fluorine', 'organic', 'organic fluorine', 'elemental fluorine', 'organic fluorine compounds', 'elemental', 'studies organic', 'studies organic fluorine', 'organic compounds', 'fluorination', 'studies', 'using elemental', 'using elemental

Topic Cluster	Number of References	Keywords/Topic Signature
		fluorine', 'synthesis', 'reactions', 'fluorine organic', 'chemistry organic', 'chemistry organic fluorine'
11	**764**	**'calcium', 'sodium', 'effect', 'effect sodium', 'phosphate', 'crystals', 'calcium phosphate', 'growth', 'effects', 'acid', 'monofluorophosphate', 'solution', 'study', 'bone', 'sodium monofluorophosphate', 'formation', 'calcium crystals', 'potassium', 'kinetics', 'effect calcium'**
12	400	'electron', 'barium', 'fluorine', 'electron transfer', 'beam', 'electron beam', 'resonance', 'transfer', 'electron paramagnetic', 'paramagnetic', 'crystals', 'electron paramagnetic resonance', 'paramagnetic resonance', 'study', 'ions', 'electron diffraction', 'diffraction', 'lithium', 'spectra', 'pi'
13	**670**	**'enamel', 'dental', 'dental fluorosis', 'fluorosis', 'uptake', 'dental enamel', 'effect', 'surface', 'remineralization', 'tooth', 'plaque', 'lesions', 'surface enamel', 'concentration', 'application', 'acid', 'demineralization', 'study', 'caries', 'teeth'**
14	432	'rare', 'earth', 'rocks', 'rare earth', 'minerals', 'ree', 'rich', 'fluid', 'magmatic', 'elements', 'ore', 'fluorine', 'bearing', 'granite', 'melt', 'granites', 'deposits', 'hydrothermal', 'fluids', 'fluorite'
15	**334**	**'phase', 'gas phase', 'gas', 'fluorine', 'solid phase', 'solid', 'phase transfer', 'reactions', 'phase transitions', 'phase chronic', 'transfer', 'skeletal phase chronic', 'skeletal phase', 'phase transition', 'transitions', 'liquid', 'transition', 'synthesis', 'beta', 'containing'**
16	517	'lithium', 'lif', 'thermoluminescent', 'response', 'thermoluminescence', 'dosimetry', 'dose', 'dosimeters', 'tld', 'energy', 'radiation', 'dosimeter', 'neutron', 'thermoluminescent dosimeters', 'lithium crystals', 'crystals', 'gamma', 'thermoluminescence lithium', 'thermal', 'thermoluminescent lithium'
17	236	'fluorine chemistry', 'chemistry', 'fluorine', 'award', 'work fluorine', 'creative', 'work fluorine chemistry', 'acs', 'creative work fluorine', 'creative work', 'acs award', 'issue', 'preface', 'special issue', 'award creative work', 'award creative', 'work', 'special', 'acs award creative', 'chemistry fluorine'
18	**194**	**'fluoridated', 'fluoridated water', 'water', 'caries', 'milk', 'fluoridated milk', 'communities', 'apatites', 'non fluoridated', 'effect fluoridated', 'salt', 'fluoridated hydroxyapatites', 'hydroxyapatites', 'fluoridated salt', 'fluoridated area', 'fluoridated apatites', 'non', 'dental', 'effect', 'nonfluoridated'**
19	370	'aluminum', 'content', 'fluorine content', 'fluorine', 'water', 'teas', 'waters', 'content water', 'effect', 'cryolite', 'foods', 'bottled', 'sodium', 'kinetics', 'effects', 'content bottled', 'acid', 'sodium aluminum', 'magnesium', 'molten'
20	**158**	**'topical', 'topical application', 'application', 'topical applications', 'enamel', 'applications', 'uptake', 'topical agents', 'caries', 'effect topical', 'effect', 'topical treatment', 'agents', 'stannous', 'treatment', 'dental', 'plaque', 'following', 'solutions', 'following topical'**
21	382	'properties', 'properties fluorine', 'thermodynamic properties', 'thermodynamic', 'fluorine', 'physical', 'physical properties', 'effect', 'production streptococcus', 'production streptococcus mutans', 'structure properties', 'structure', 'streptococcus', 'sodium', 'production', 'magnetic properties', 'containing', 'properties sodium', 'preparation properties', 'influence'
22	541	'solutions', 'chloride', 'aqueous', 'aqueous solutions', 'melts', 'extraction', 'chloride melts', 'niobium', 'ammonium', 'tantalum', 'phosphate', 'exchange', 'anion', 'potassium', 'bromide', 'titanium', 'uranium', 'iv', 'sodium', 'ion'
23	559	'ion', 'reactions', 'fluorine ion', 'involving ion', 'reactions involving ion', 'reactions involving', 'fluorine', 'involving', 'ion induced', 'presence ion', 'induced', 'ions', 'effect', 'ion implantation', 'presence', 'implantation', 'study', 'exchange', 'effect ion', 'ion exchange'

Topic Cluster	Number of References	Keywords/Topic Signature
24	156	'graphite', 'intercalation', 'graphite intercalation', 'intercalation compounds', 'compounds', 'fluorine', 'graphite intercalation compounds', 'intercalated', 'intercalation compound', 'preparation', 'fluorine graphite', 'graphite intercalation compound', 'intercalated graphite', 'fluorine intercalated graphite', 'fluorine intercalated', 'compound', 'graphite compounds', 'conductivity', 'electrical', 'preparation graphite'
25	**9,450**	**'effect', 'synthesis', 'new', 'reaction', 'using', 'study', 'effects', 'polyvinylidene', 'structure', 'compounds', 'studies', 'reactions', 'activity', 'formation', 'containing', 'analysis', 'based', 'acid', 'high', 'use'**

Bold font indicates topic clusters that appear to include studies that meet the populations, exposures, comparators, and outcomes (PECO) criteria based on the listed terms or included studies identified in the crosswalk with published assessments (N = 12,666 studies).

Table C-7. Targeted Topic Extraction Results for Fluoride Literature with No Tag in SWIFT-Review from Cluster 25 of Targeted Topic Extraction of Clusters 3, 12, and 14

Topic Cluster	Number of References	Keywords/Topic Signature
1	140	['stability', 'chromyl', 'thermal', 'thermal stability', 'fluorine', 'chemistry chromyl', 'containing', 'formation', 'synthesis', 'apatites', 'metal stability', 'effect', 'cro2f2', 'stability toothpastes', 'concentration stability', 'chemistry', 'stability chromyl', 'influence', 'study', 'new']
2	**444**	**['mutans', 'streptococcus', 'streptococcus mutans', 'formation', 'glass', 'production', 'ph', 'development', 'results', 'study', 'caries', 'interaction', 'effect', 'action', 'biofilm', 'acid', 'cariogenic', 'release', 'inhibition', 'old']**
3	431	['reaction', 'methyl', 'solvent', 'solvents', 'mediated', 'tetramethylammonium', 'organic', 'effect', 'potassium', 'reactions', 'organic solvents', 'fluorine', 'using', 'conditions', 'products', 'proton', 'study', 'formation', 'substitution', 'basic']
4	427	['chemistry', 'exchange', 'using', 'vi', 'sulfonyl', 'free', 'silica', 'sufex', 'salts', 'conditions', 'click', 'reagents', 'new', 'reaction', 'reactions', 'synthesis', 'sulfur', 'syntheses', 'substituted', 'catalysts']
5	256	['gas', 'sulfur', 'carbon', 'diffusion', 'mixtures', 'dioxide', 'sulfur hexafluoride', 'hexafluoride', 'emissions', 'liquid', 'gases', 'sf6', 'sulfur dioxide', 'reaction', 'gas liquid', 'chromatography', 'carbon dioxide', 'analysis', 'greenhouse', 'using']
6	**355**	**['effect', 'hydroxyapatite', 'growth', 'kinetics', 'crystal', 'bone', 'uptake', 'apatite', 'crystal growth', 'stannous', 'crystals', 'plaque', 'uptake hydroxyapatite', 'activity', 'surface', 'ph', 'dissolution', 'dentifrice', 'mechanism', 'salivary']**
7	**520**	**['studies', 'enzyme', 'mechanism', 'binding', 'site', 'substrate', 'glucose', 'enzymes', 'active', 'molecular', 'inhibition', 'active site', 'experimental', 'inhibitors', 'oxidase', 'cytochrome', 'activity', 'protein', 'irreversible', 'reaction']**
8	357	['infrared', 'based', 'spectroscopy', 'detection', 'pressure', 'strength', 'trace', 'analysis', 'study', 'transitions', 'double', 'using', 'atmospheric', 'ir', 'identification', 'methyl', 'used', 'power', 'spectroscopic', 'fluorine']
9	319	['systems', 'spectra', 'containing', 'vibrational', 'raman', 'vibrational spectra', 'infrared', 'silicate', 'raman spectra', 'studies', 'study', 'evaluation', 'effect', 'synthesis', 'infrared spectra', 'new', 'spectroscopy', 'temperature', 'formation', 'water']
10	147	['characterization', 'applications', 'film', 'synthesis', 'synthesis characterization', 'polyvinylidene', 'oriented', 'polyvinylidene film', 'new', 'transducer', 'sensor', 'preparation characterization', 'preparation', 'talk applications', 'films', 'poly', 'polyvinyl', 'talk', 'n3nfo', 'chemical']
11	**305**	**['effects', 'oral', 'health', 'oral health', 'oral bacteria', 'dental', 'hygiene', 'stannous', 'bacteria', 'public', 'oral hygiene', 'plaque', 'public health', 'streptococci', 'use', 'oral streptococci', 'effect', 'exposure', 'care', 'dental health']**

Topic Cluster	Number of References	Keywords/Topic Signature
12	165	['polyvinylidene', 'pyroelectricity', 'pyroelectricity polyvinylidene', 'piezoelectricity', 'piezoelectricity pyroelectricity polyvinylidene', 'piezoelectricity pyroelectricity', 'films', 'polyvinylidene films', 'pyroelectric', 'hysteresis', 'effect polyvinylidene', 'crystallization polyvinylidene', 'pvf2', 'transducers', 'crystallization', 'corona', 'polyvinylidene pvf2', 'effect', 'currents polyvinylidene', 'polarization']
13	**720**	**['reactions', 'new', 'compounds', 'activity', 'group', 'fluorine', 'containing', 'synthesis', 'rights reserved', 'rights', 'reserved', 'derivatives', 'science', 'inhibitors', 'active', 'fluorine containing', 'anions', 'reaction', 'synthesized', 'structure']**
14	449	['magnesium', 'oxide', 'preparation', 'compounds', 'technique', 'low', 'adsorption', 'new', 'potassium', 'using', 'reaction', 'structure', 'synthesis', 'method', 'high', 'use', 'temperature', 'strength', 'windows', 'study']
15	247	['synthesis', 'ionic', 'ionic conductivity', 'conductivity', 'liquids', '10', 'article', 'potassium', 'nanoparticles', 'ionic liquids', 'carbonate', 'metal', 'butyl', 'doi', 'doi 10', 'new', 'precursors', 'using', 'upconverting', 'upconverting nanoparticles']
16	**88**	**['silver', 'diamine', 'diammine', 'silver diamine', 'ag', 'diammine silver', 'reaction', 'diffusion silver', 'polymorphism silver', 'structure silver', 'diamine diammine', 'reaction silver', 'ii', 'teeth', 'polymorphism', 'silver nitrate', 'dental', 'sdf', 'diamine silver', 'detection formation']**
17	**109**	**['addition', 'sealants', 'fissure', 'fissure sealants', 'pit fissure', 'pit', 'release', 'pit fissure sealants', 'strength', 'sealant', 'michael', 'effect', 'michael addition', 'bond', 'varnish', 'radical', 'molars', 'permanent molars', 'unsaturated', 'addition bromine']**
18	366	['fluorination', 'recent', 'review', 'bonds', 'selective', 'fluorine', 'synthesis', 'years', 'strategies', 'reactivity', 'organic', 'advances', 'synthetic', 'recent advances', 'applications', 'sources', 'reactions', 'electrophilic', 'direct', 'molecules']
19	**169**	**['release', 'protease', 'enzyme', 'serine', 'purified', 'activity', 'inhibited', 'ph', 'phenylmethylsulfonyl', 'purification', 'kda', 'gel', 'orthodontic', 'serine protease', 'extracellular', 'degrees', 'inhibitor', 'bonding', 'chromatography', 'molecular']**
20	257	['acids', 'lewis', 'acid', 'lewis acids', 'lewis acid', 'difference', 'acidity', 'lewis acidity', 'reaction', 'formation', 'boron', 'carboxylic acids', 'carboxylic', 'bidentate', 'amino acids', 'amino', 'synthesis', 'acidic', 'based', 'fluorinated']
21	486	['analysis', 'metal', 'structure', 'interactions', 'role', 'function', 'coordination', 'compounds', 'fluorine', 'chemistry', 'new', 'alkaline', 'developments analysis', 'synthesis', 'alkali', 'alkali metal', 'transition', 'using', 'molecular', 'reactions']
22	**2,000**	**['reply', 'study', 'toothpaste', 'intake', 'bone', 'fluoridation', 'use', 'toxicity', 'treatment', 'concentration', 'teeth', 'uptake', 'dentifrices', 'levels', 'plaque', 'toothpastes', 'dentifrice', 'potassium', 'evaluation', 'concentrations']**
23	50	['research', 'international', 'society', 'society research', 'international society', 'conference', 'international society research', 'conference international society', 'conference international', 'abstracts', 'papers', 'abstracts papers', 'papers presented', 'september', 'abstracts papers presented', 'recent', 'international research', 'presented', 'report', '2018']
24	**341**	**['metabolism', 'alpha', 'high', 'mass', 'beta', '15', 'synthesis', 'chain', '14', 'derivatives', '13', 'incorporation', 'nmr', 'analyses', 'deuterium', 'prepared', 'activities', '19', 'reaction', '23']**
25	**302**	**['materials', 'dentin', 'acid', 'potential', 'restorative materials', 'restorative', 'form', 'resistant', 'containing', 'layer', 'release', 'treatment', 'effect', 'releasing', 'treated', 'resistance', 'surface', 'root', 'root dentin', 'synthesis']**

Bold font indicates topic clusters that appear to include studies that meet the populations, exposures, comparators, and outcomes (PECO) criteria based on the listed terms or included studies identified in the crosswalk with published assessments (N = 5,353 studies).

C.2.2. Full-Text Prioritization

Prior to full-text review, the 6,562 English language studies that advanced from title-abstract screening were subjected to another round of prioritization to identify studies that would be most informative for dental fluorosis and neurodevelopmental outcomes. After review and deduplication of the studies from the crosswalk of existing assessments, 181 studies were advanced to full-text screening.

C.2.2.1. SWIFT-Review Filtering

SWIFT-Review filtering was used to more efficiently identify studies informing dental fluorosis and neurodevelopmental effects. The EPA developed and validated a custom search string (Table C-8). All other filters were the defaults available in SWIFT-Review.

Table C-8. Custom SWIFT-Review Search String for Dental and Neurodevelopmental Studies

SWIFT-Review Filter Name	SWIFT-Review Filter Search String
Dental	journal: ("BMC Oral Health" OR "Caries Research" OR "Journal of Dentistry" OR "Journal of Dental Research") OR mesh_mh("Dental caries" OR "Fluorosis, Dental" OR Dentistry OR "Dental health services" OR "Dental care") OR tiab: ("Dental caries" OR "Dental Cavities" OR "Dental fluorosis" OR "Enamel demineralization" OR "Fluoride intake" OR "Occlusal caries" OR "Interproximal caries" OR "Tooth cavities" OR "Tooth decay" OR "Tooth demineralization" OR ((tooth OR teeth) AND enamel) OR "Dental health" OR "dental pulp" OR remineralization OR "dental hard tissues" OR dentin OR "dental pulp" OR dentinogenesis OR hypomineralized OR "dentin hypersensitivity" OR "fluoride gel" OR "fluoride varnish" OR "dentin structure" OR dentition* OR xerostomia OR gingivitis OR "gum disease" OR "periodontal disease" OR periodontitis OR crown* OR "Dental implant*" OR "dental bridge*" OR "root canal*" OR "teeth braces" OR odontoblast*)
Neurodevelopment	mesh_mh: ("Attention Deficit and Disruptive Behavior Disorders" OR "Neurodevelopmental Disorders" OR "executive function" OR "processing speed" OR "dopamine" OR "serotonin" OR "glutamic acid" OR "gamma aminobutyric acid" OR "epinephrine" OR "norepinephrine" OR "glycine" OR "acetylcholine" OR glutamates) OR ((mesh_mh:(prosencephalon) OR tiab:(prosencephalon OR forebrain)) AND (mesh_code:(abnormalities.*) OR tiab:(abnormalities OR anomalies))) OR (tiab:perinatal AND (mesh_mh:("brain injuries" OR tiab:((brain AND injuries) OR (brain AND injury))))) OR tiab: (“Internalizing behavior*” OR “Internalizing behaviour*” OR “externalizing behavior*” OR “externalizing behaviour*” OR “Bayley Scale*” OR “McCarthy Scale*” OR “oppositional behavior*” OR “oppositional behaviour*” OR “executive function*” OR “processing speed*” OR "visuospatial " OR “behavior disorder”~3 OR “behavior disordered”~3 OR “behavior disorders”~3 OR “behaviour disorder”~3 OR “behaviour disordered”~3 OR “behaviour disorders”~3 OR “development disorder”~3 OR “development disorders”~3 OR “developmental disorder”~3 OR “developmental disorders”~3 OR “language disorder”~3 OR “language disorders”~3 OR “speech disorder”~3 OR “speech disorders”~3 OR “asd” OR "conduct disorder*" OR "intellectual disabilit*" OR "learning deficit*" OR "learning disabl*" OR "learning disorder*" OR "learning problem*" OR “neuropsychological” OR “oppositional defiant” OR "anxiety-like behavior*" OR "anxious behavior*" OR "associative learn*" OR "behavioral abnorm*" OR "behavioral alteration*" OR "behavioral assay*" OR "Behavioral change*" OR "behavioral deficit*" OR "behavioral develop*" OR "behavioral diff*" OR "Behavioral effect*" OR "behavioral respons*" OR "behavioral stud*" OR "behavioral technique*" OR "behavioral test*" OR "Behavioural

SWIFT-Review Filter Name	SWIFT-Review Filter Search String
	change*" OR "cognitive abilit*" OR "cognitive alter*" OR "cognitive assess*" OR "cognitive defic*" OR "cognitive impair*" OR "cognitive problem*" OR "cognitive respons*" OR "intellectual abilit*" OR "intellectual develop*" OR "intelligence scale*" OR "intelligence score*" OR "intelligence test*" OR "language comprehen*" OR "language outcome*" OR "learning abilit*" OR "learning capabilit*" OR "learning error*" OR "memory abilit*" OR "memory consolidat*" OR "memory dysfunction*" OR "memory error*" OR "memory function*" OR "memory impair*" OR "memory observation*" OR "memory skill*" OR "open-field behavior*" OR "spontaneous behavior*" OR “ability to learn” OR “behavioral teratology” OR “behavioral toxicity” OR “borderline intelligence” OR “Brain Development” OR “brain learning” OR “child intelligence” OR “children's intelligence” OR “children's IQ” OR “child's intelligence” OR “child's IQ” OR “cognitive development” OR “cognitive dysfunction” OR “cognitive function” OR “cognitive outcome” OR “cognitive performance” OR “decreased learning” OR “depressive behavior” OR “depressive-like behavior” OR “emotional behavior” OR “excellent intelligence” OR “full-scale intelligence” OR “general cognitive index” OR “impaired learning” OR “impairment of learning” OR “impairment of memory” OR “infant intelligence” OR “intellectual impairment” OR “intellectual performance” OR “intelligence development” OR “intelligence evaluation” OR “intelligence in children” OR “intelligence level” OR “intelligence loss” OR “intelligence quotient” OR “intelligence response” OR “learning and memory” OR “learning impairment” OR “learning observation” OR “learning-memory” OR “low intelligence” OR “lower intelligence” OR “Maze*” OR “memory retention” OR “mobility index” OR “motor behavior” OR “motor coordination” OR “Neurofunction*” OR “object recognition memory” OR “performance behavior” OR “reference memory” OR “retention memory” OR “scale of intelligence” OR “spatial learning” OR “spatial memory” OR “Wechsler Scale” OR “working memory” OR “dopamin*” OR “seroton*” OR “5-hydroxytryptamine” OR “5-HT” OR “glutamate*” OR “gamma aminobutyric acid*” OR “GABA” OR “epinephrine” OR “norepinephrine” OR “glycine” OR “acetylcholine” OR “glutamic acid*” (dandy AND walker AND syndrome) OR ("dandy walker" AND syndrome) OR (neural AND tube AND defect) OR (neural AND tube AND defects) OR ("histopath*" AND (brain OR hippocampus OR "gray matter" OR "white matter" OR cerebellum OR cerebral OR cerebrum OR cortex OR cerebellar OR brainstem or "brain stem")) OR ("histol*" AND (brain OR hippocampus OR "gray matter" OR "white matter" OR cerebellum OR cerebral OR cerebrum OR cortex OR cerebellar OR brainstem or "brain stem"))) OR tiab_stemmed: ("adhd" OR "adolesc attent" OR "adolesc bipolar" OR "adolesc dissoci" OR "adolesc major depress" OR "adolesc schizophrenia" OR "affect disord" OR "assess cognit" OR "attent deficit" OR "autism" OR "autist" OR "base cognit behavior" OR "behavior problem" OR "behavior problem children" OR "chang cognit" OR "child abus dissoci" OR "child develop" OR "childhood anxiet" OR "childhood trauma dissoci" OR "children attent" OR "children behavior" OR "children development coordin" OR "children development disabl" OR "children high function" OR "children learn disabl" OR "children obsess compuls" OR "children pervas development" OR "children read difficult" OR "comorbid attent deficit" OR "compuls" OR "conduct disord" OR "coordin disord" OR "dandy walker syndrome" OR "de lang syndrom" OR "debrief after psycholog" OR "deficit children" OR "deficit disord" OR "deficit hyperact" OR "development coordin" OR "development dela" OR "development disabl" OR "development disord" OR "development screen" OR "diagnos mental retard" OR "disabl children" OR "disord adolesc girl" OR "disrupt behavior" OR "dissoci childhood" OR "handicap children" OR "hydrocephalus" OR "hydromyelia" OR "hyperact disord" OR "individu development disabl" OR "intellectu disabl" OR "lead poison children" OR "learn difficult" OR "learn disabl" OR "learn disord" OR "learn impair" OR "lesch nyhan syndrom" OR "matern phenylketonuria" OR "mental defect" OR "mental defici" OR "mental disabl" OR "mental handicap" OR "mental retard" OR "mental subnorm" OR "motor develop" OR "neural tube defect" OR "neurodevelopment" OR "normal pressur hydrocephalu" OR "pervas development" OR "poor reader" OR "premenstru dysphor" OR "problem behavior" OR

SWIFT-Review Filter Name	SWIFT-Review Filter Search String
	"problem children" OR "problem mental" OR "profoundl retard" OR "psychomotor develop" OR "read difficult" OR "read disabl" OR "retard adolesc" OR "retard child" OR "retard children" OR "retard patient" OR "retard person" OR "retard syndrom" OR "sever retard" OR "Sjogren-Larsson syndrome" OR "special educ" OR "special school" OR "spectrum disord" OR "strength difficulti questionnair" OR "studi attent deficit" OR "subnorm children" OR "x link mental" OR "x syndrom")

mesh_mh = MeSH term; tiab = title-abstract.

Following the development and validation of the dental and neurodevelopmental strings, all studies included during title-abstract screening (n = 6,562) underwent SWIFT-Review filtering for health outcome and evidence stream tagging (Table C-9).

Table C-9. SWIFT-Review Filtering Results for Health Outcome and Evidence Stream

	Evidence Stream[a,b]			
Health Outcome[c]	**Human**	**Animal**	**No Evidence Stream Tag**	**Total**
Dental	3,734 (76%)	1,527 (63%)	207 (37%)	4,374 (67%)
Neurodevelopment	205 (4%)	164 (7%)	4 (1%)	297 (5%)
Cancer	144 (3%)	108 (4%)	9 (2%)	217 (3%)
No Priority Health Outcome Tag	1,358 (27%)	1,018 (42%)	346 (63%)	2,398 (37%)
Total	4,913 (75%)	2,435 (37%)	551 (8%)	6,562 (100%)

[a]Studies could receive more than one health outcome or evidence stream tag, which might lead to percentages not adding to 100%.
[b]The percentage of studies with each evidence stream tag was determined from the overall total number of references tagged in Swift-Review (n = 6,562).
[c]The percentage of studies within each health outcome category was determined by the number of studies tagged to each specific evidence stream noted in the column title.

Given the large number of studies tagged to both human tags ("Human" or "Epidemiologic Quantitative Analyses" from Section 4.2.3.1) and "Dental" tags, studies tagged to both the "Epidemiologic Quantitative Analyses" and "Dental" tags were moved to full-text review since the toxicity assessment is scoped to dose-response analysis and the "Epidemiological Quantitative Analysis" tagged studies are most likely to be relevant for dose-response analysis. A subset of studies received a priority outcome tag but did not receive an evidence stream tag. Titles for these studies were reviewed by at least two scientists and a group of studies with potentially relevant information were grouped as "Select No Tag Priority Studies" and sent for screening.

After removing the 181 studies identified from other assessments already forwarded for full-text review, studies from the following groups were prioritized. Total counts are presented for each group, and some studies are represented in more than one group.

- Human – Neurodevelopmental (n = 117)
- Human – Cancer (n = 141)
- Epidemiologic Quantitative Analyses – Dental (n = 2,259)
- Animal – Neurodevelopmental (n = 142)
- Animal – Cancer (n = 106)
- Select No Tag Priority Studies (n = 40)

In addition, 113 non-English language studies were forwarded for full-text screening (Section 4.2.6, Non-English Language Studies provides additional details on the approach for these studies).

C.2.2.2. Supervised Clustering of Dental Studies

After manually screening the full texts of over 1,000 studies in the Epidemiologic Quantitative Analyses – Dental category and identifying only 25 relevant dental studies with potentially sufficient quantitative information for dose-response analysis, supervised clustering was conducted to 1) prioritize the remaining studies in the Epidemiologic Quantitative Analyses – Dental category to identify those most likely to contain adequate data for dose-response analysis, and 2) prioritize all studies in the Human – Dental category to identify those most likely to be relevant and provide data for dose-response analyses.

The 25 Epidemiologic Quantitative Analyses – Dental studies identified as relevant and potentially containing sufficient quantitative data to support dose-response analysis during initial full-text screening were used as positive seed studies to predict the remaining dental studies most likely to be relevant. First, the remaining studies tagged to Epidemiologic Quantitative Analyses – Dental that had not been screened at the full-text level (n = 712) underwent supervised clustering using the 25 seed studies. Results are provided in Table C-10. Then, studies that received ensemble scores 1–6 moved forward for full-text screening and literature inventory (n = 193 unique new studies), bringing the total number of Epidemiologic Quantitative Analyses – Dental studies for full-text screening to 1,701.

Table C-10. Results from Supervised Clustering of Epidemiologic Quantitative Analyses – Dental Studies Awaiting Full-Text Screening

Ensemble Scores	Number of Studies	Number of Seed Studies
0	519	0
1	72	1
2	33	2
3	29	0
4	35	1
5	32	9
6	17	12
Grand Total	**737**	**25**

Additionally, all unique studies tagged to Human – Dental (i.e., those that did not also receive the Epidemiologic Quantitative Analyses tag) underwent supervised clustering using the same 25 dental seed studies. None of these 1,306 studies had been screened at the full-text level. Results are in Table C-11. Studies that received ensemble scores 1–6 moved forward for full-text screening and literature inventory (n = 317 unique new studies).

Table C-11. Results from Supervised Clustering of Human – Dental Studies Awaiting Full-Text Screening

Ensemble Scores	Number of Studies	Number of Seed Studies
0	989	0
1	178	1
2	71	0
3	33	2
4	18	8
5	24	10
6	18	4
Grand Total	**1,331**	**25**

C.3. Literature Screening Strategies for ADME Studies

The EPA is also reviewing the body of literature to identify studies containing fluoride absorption, distribution, metabolism, and excretion (ADME) information. The methods used to identify ADME studies are similar to those described in Section 4.2, with modifications tailored to this specific evidence type.

C.3.1. PECO Criteria for ADME Literature Survey

Separate PECO criteria (Table C-12) were developed to guide the review of relevant ADME studies and specify whether these studies would be included in the ADME literature survey. During title-abstract screening, physiologically based pharmacokinetic (PBPK) models are included so that these studies could be leveraged during future steps of the review (e.g., during pharmacokinetic (PK) modeling evaluation). Additionally, review articles or studies written in a non-English language are included during title-abstract ADME screening, as noted in the PECO criteria.

Table C-12. PECO Criteria for Screening ADME Studies

PECO Element	Inclusion Criteria
Population	**Human**: Any population and lifestage (occupational or general population, including children and other sensitive populations): whole organism, tissues, individual cells, or biomolecules.
Exposure	**Human:** Any exposure to the chemicals listed via *oral routes, any injection-type routes (*e.g., *intravenous, sub-cutaneous, intraperitoneal, intramuscular)*, or *unknown/multiple routes* (e.g., biomonitoring studies). **Exposure to fluoride, and/or related forms, including but not limited to:** fluoride (*CASRN 16984-48-8*). *Other names/forms*: ammonium fluoride calcium fluoride disodium hexafluorosilicate fluoride ion fluorine fluorosilic acid hydrofluoric acid*, ion(1-) hydrofluorosilicate

PECO Element	Inclusion Criteria
	hydrofluosilicic acid perfluoride potassium fluoride sodium hexafluorosilicate sodium fluoride (natrium fluoride) sodium fluorosilicate sodium silicofluoride amine fluoride olaflur sodium monofluorophosphate stannous fluoride amine fluoride acidulated phosphate fluoride The following compounds will be excluded: aluminum fluoride (due to complex interactions with fluoride) PFAS and fluorotelomers (even if fluoride is a measured metabolite) fluoridated organic compounds, e.g., fluoridated pyrimidines (even if fluoride is a measured metabolite) fluoride metal salts not listed earlier, such as: silver diamine fluoride titanium tetrafluoride bismuth fluoride *…OR…* Any exposure to the chemicals listed via inhalation representing studies with fluoridated anesthetics if they provide a concentration of the fluoride ion as a metabolite. isoflurane methoxyflurane enflurane sevoflurane halothane desflurane
Comparators	Multiple measurements to relate metrics of exposure or evaluate ADME properties. Examples include estimated/measured exposure and some biomarker concentration (e.g., urine), measurements of multiple biomarkers (e.g., urine, plasma, and/or saliva), or measurements of a single biomarker over time (e.g., urine at multiple times during pregnancy). Estimated exposure would include an estimated drinking water concentration for that cohort. *A single biomarker measure (e.g., only urine or only plasma) at one time, without a measurement or estimate of exposure would be excluded, unless it contains other information on ADME properties*.
Outcomes	Any examination of fluoride 1) absorption of dose through gastrointestinal tract or injection-type route, (2) distribution across biological compartments (e.g., blood, tissue, excreta), and/or (3) excretion. Studies describing pharmacokinetic (PK) models for fluoride will be included. **NOTE**: Fluoride, being the ion of the element fluorine, is NOT metabolized. **NOTE**: Relevant reviews and foreign language studies will be included at title-abstract screening.

PECO = populations, exposures, comparators, and outcomes; ADME = absorption, distribution, metabolism, and excretion; CASRN = Chemical Abstracts Service registry number; PFAS = per- and polyfluoroalkyl substances; PK = pharmacokinetic.
*Dosing with hydrofluoric acid could include occupational exposures to hydrofluoric acid (i.e., poisoning), which would not be relevant to ADME.

C.3.2. Prioritization Prior to Title-Abstract Screening

Prior to ADME title-abstract screening, prioritization of studies is conducted. The prioritization approaches largely follow the methods used for health effects studies, described in Section 4.2. All studies identified in the literature search are imported into SWIFT-Review software for filtering (Section 4.2.3.1). Studies that receive either the Human or the Epidemiologic Quantitative Analyses evidence stream tag and an ADME health outcome tag are advanced for additional prioritization via topic extraction (Section 4.2.3.3).

C.3.3. Title-Abstract Screening

Screening of the title and abstract of each study against the ADME PECO criteria is performed by two independent reviewers per study using Sciome's SWIFT-Active Screener software. If results from the two independent screeners conflict, a senior-level tertiary screener resolves the conflict. Studies determined to meet the ADME PECO criteria are designated as 'included' and advance to full-text review. SWIFT-Active Screener has a machine-learning component that is used to predict relevant references and expedite the screening of large evidence bases by prioritizing the references predicted to be relevant. Studies are screened until the software's algorithm predicts that 95% of the relevant studies have been screened and included (95% threshold).

Following completion of title-abstract ADME screening, additional review by a senior-level screener is conducted according to the procedure outlined in Section 4.2.4.

C.3.4. Prioritization Prior to Full-Text Review

Dental-specific ADME studies identified during title-abstract screening are subject to further prioritization prior to full-text review. While many dental-specific studies provide useful ADME information related to fluoride uptake in saliva and teeth, many focus on fluoride's role in cavity prevention or interaction with the oral microbiome. Therefore, it is important to determine which dental-specific ADME studies inform the systemic uptake, distribution, and excretion of fluoride in the body and are relevant to the future Fluoride Human Health Toxicity Assessment. Dental-specific ADME studies identified during title-abstract screening undergo additional SWIFT-Review filtering using a modified version of the 'dental' tag is used to tag health effects studies (Section 4.2.3.1 describes how SWIFT-Review is used for prioritization and Table C-8 provides the dental-specific search string), to help identify studies likely to contain the dental-specific ADME information. To determine dental studies relevant to the toxicity assessment's ADME, studies assigned this dental ADME tag undergo additional review to determine their suitability for advancement to full-text review using the same tools outlined in Section 4.2.3.2 and 4.2.3.3 as needed. ADME studies that do not receive this dental tag are advanced to full-text review.

C.3.5. Full-Text Review and Literature Inventory Development

Full-text ADME screening and literature inventory steps are completed concurrently using ICF's Litstream™ software. Studies deemed PECO-relevant during ADME title-abstract screening or identified as containing ADME information during full-text review of health effects studies undergo an additional step during which key ADME data from the studies (e.g., stage of ADME evaluated, chemical characterization, study design details, and measurement details of fluoride in

media and biological matrices) are extracted into Litstream™ to create a literature inventory. A primary reviewer completes the initial screening (and literature inventory extraction if applicable) and all studies undergo secondary review for quality assurance by a senior-level reviewer. Studies that are not PECO-relevant according to the ADME PECO criteria are briefly tagged but do not undergo data extraction for the literature inventory.

C.3.6. ADME Studies Cited in Existing Peer-Reviewed Sources

In addition to the crosswalks identified in Section 4.1.2, the EPA is reviewing the following publications to identify cited studies containing ADME information to supplement the database search results and target potential data gaps:

- The use of urinary fluoride excretion to facilitate monitoring fluoride intake: A systematic scoping review (Idowu et al., 2019)
- Updated consumer risk assessment of fluoride in food and drinking water including the contribution from other sources of oral exposure (EFSA Scientific Committee, 2025)
- Fluoride Metabolism in Pregnant Women: A Narrative Review of the Literature. (Castiblanco-Rubio and Martinez-Mier, 2022) https://doi.org/10.3390/metabo12040324

C.4. Additional Pharmacokinetic Models

Whereas PK models describe the systemic uptake and distribution of fluoride, tissue-specific models offer mechanistic insights into individual organ systems. These tissue-specific models are designed to quantify fluoride ADME within particular regions or organ systems. Although they cannot characterize the entire ADME process from exposure to elimination independently, they provide valuable information regarding distinct components of fluoride's biological pathway.

C.4.1. Bone Uptake

As described in the PK key science issue (Section 2.4), fluoride preferentially distributes to the skeleton in an age-dependent manner, which results in bone acting as a long-term storage site that slowly releases fluoride back into the bloodstream. Because urine fluoride concentration measurements reflect recent fluoride intake and bone releases fluoride slowly, accounting for bone uptake and release could be necessary when using blood and urine fluoride measurements to gauge long-term exposure. Therefore, models that include bone uptake and release aid in interpretation of fluoride biomonitoring data derived from long-term exposure rather than just recent intake.

For fluoride, the model structure outlined by Hawkins et al. (1992) is extensively used to assess skeletal fluoride ion kinetics in positron emission tomography (PET) imaging. Their three-compartment model includes vascular (plasma), extravascular (bone extracellular fluid), and bone-mineral ("bound") compartments, with rate constants for plasma-tissue exchange and for mineral incorporation and release. Using this three-compartment model structure and parameters estimated in healthy adults, Hawkins et al. (1992) showed that the model accurately reflects the net transfer of fluoride from plasma into bone mineral. They also found that release of fluoride from bone back to plasma is slow (on the order of days). Together, these features provide a

mechanistic basis for using bone-informed kinetics to interpret how bone influences blood concentrations of fluoride and to infer longer-term exposure from biomonitoring data.

Subsequent applications of the Hawkins et al. (1992) model have shown how fluoride uptake varies across skeletal sites and have illustrated how the model can describe uptake into newly formed bone. Frost et al. (2009) estimated rate constants across the spine and humerus to show that the rate of uptake into the bone reflected osteoblastic activity. The same bone-focused framework is also applicable to characterizing fluoride uptake in newly formed bones. de Ruiter et al. (2024) showed that fluoride uptake into newly formed bone can reflect ongoing bone formation and remodeling over time, supporting its use as a marker of new skeletal tissue formation. Overall, bone-specific models built on the Hawkins et al. (1992) framework show that fluoride uptake differs by skeletal site and by bone-forming activity. Incorporating bone uptake and release into pharmacokinetic models would help to ensure that bone is properly accounted for when estimating lifestage-specific internal fluoride levels. However, if bone release of fluoride is never more than a small fraction of uptake (e.g., less than 10%), then accounting for release might not significantly improve model predictions while complicating the model analysis.

C.4.2. Oral Uptake

The fluoride exposures considered for the toxicity assessment occur via the oral route through drinking water, foods, dietary supplements or medications, and dental products (e.g., dentifrices, mouth rinses). Regardless of source, the dose first enters the oral environment, where it mixes with saliva, contacts enamel and biofilms, and interacts with oral micro-reservoirs such as soft tissues, plaque, and stagnation zones (between teeth and under the tongue). These within-mouth processes produce rapid post-ingestion or post-application changes in salivary fluoride, followed by a slower release from the micro-reservoirs. This slower release can sustain low salivary fluoride for hours and repeated daily exposures elevate the between-exposure baseline concentration of fluoride in the saliva. Representing oral pharmacokinetics is essential for interpreting salivary/plaque measurements and is likely needed to reconcile some study-specific exposures (e.g., drinking water concentration, diet, and brushing frequency) with measurable biomarkers (e.g., serum or urine fluoride concentrations).

To quantitate salivary changes in fluoride concentration following application of toothpaste, Duckworth and Morgan (1991) developed a pharmacokinetic model to predict salivary concentrations following oral exposure. This two-compartment salivary model represents the fast clearance from the saliva and the slow reservoir-mediated release from oral tissues/plaque that can generalize beyond dentifrice application to other oral exposure sources (e.g., drinking water, food). The rate of salivary transfer to the stomach in the model provides a means to interface oral PK with systemic PK to aid in translated external exposures to relevant biomarkers. The salivary kinetics and uptake in oral tissues likely impact systemic bioavailability in the short-term (minutes or early hours) following an oral exposure but might have minimal effect on total absorption. The impact of including oral pharmacokinetics on predicting lifestage- and endpoint-specific biomarkers will be further evaluated for use in the toxicity assessment.

Appendix D. Chemical Interference Effects on Fluoride Quantitation Using Zirconium-Alizarin Colorimetric/Photometric Methods

Table D-1 outlines chemical interferences on fluoride measurement when using zirconium-alizarin colorimetric/photometric methods. The assessment will consider the presence and concentration of known interferents when evaluating exposure measurement in studies using this analytical chemistry method. Examples of specific limits to consider can be found in the references below.

Table D-1. Chemical Interferences on Fluoride Measurement When Using Zirconium-Alizarin Colorimetric/Photometric Methods

Analytical Method	Chemical Interference	Effect on F^- Measurement	Interferent Concentration
Colorimetric/Photometric (Zirconium-Alizarin reagent)	Phosphate (PO_4^{3-})	Underestimates/Overestimates	Quantitative tolerance limits can vary with experimental conditions, and negligible interference is observed only when the interferent concentration remains below these limits.[a]
	Sulfate (SO_4^{2-})	Overestimates	
	Chloride (Cl^-)	Overestimates	
	Chlorine (Cl_2)	Overestimates	
	Carbonate/Bicarbonate (Alkalinity)	Underestimates	
	Calcium (Ca^{2+})	Underestimates	
	Iron (Fe^{2+}, Fe^{3+})	Underestimates	
	Aluminum (Al^{3+})	Underestimates	
	Magnesium (Mg^{2+})	Underestimates	
	Manganese (Mn^{2+})	Overestimates	

[a]Examples include (Lamar and Drake, 1955; Megregian and Maier, 1952; Taras et al., 1950; American Public Health Association, 1946; Lamar, 1945; Elvove and Goudey, 1941; Lamar and Seegmiller, 1941; Scott, 1941; Walker and Finlay, 1940; Sanchis, 1936, 1934).

Appendix E.Benchmark Dose Modeling – Hybrid Approach

The hybrid approach is a framework used in benchmark dose (BMD) modeling to convert continuous data (e.g., intelligence quotient [IQ] score and birth weights) into a probability of an adverse response by defining an abnormal cutoff (X_0) in unexposed population distribution. As illustrated in Figure E-1, the distribution of IQ response for an unexposed population is shown on the y-axis where dose equals zero ($d = 0$), and $P_0 = \Pr(X < X_0 | d = 0)$ is the probability of an unexposed population falling into the abnormal range. As exposure increases, the distribution shifts. The BMD is defined as the specific dose at which the probability of the population falling into the abnormal range equals $P_d = \Pr(X < X_0 | d = BMD)$. P_d can be defined by specifying the benchmark response (BMR) as either extra risk

$$Extra\ Risk = \frac{P_d - P_0}{1 - P_0}$$

or added risk

$$Added\ Risk = P_d - P_0\ .$$

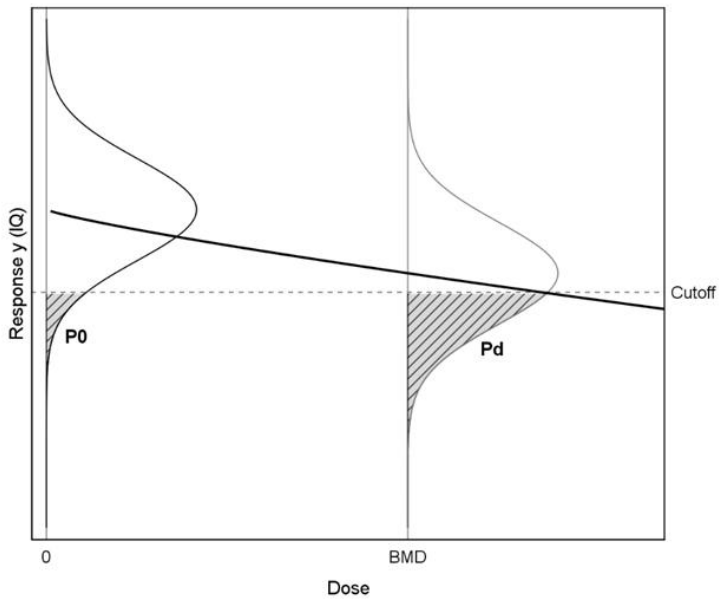


Figure E-1. Illustration of Dose-Response Modeling Using Hybrid Approach for BMR

This approach can be used for observational data where studies carry out their own dose-response modeling. The benchmark dose lower confidence limit (BMDL) is the lower bound 95% confidence limit of BMD.

Appendix F. Pharmacokinetic Model Evaluation Criteria

Physiologically based pharmacokinetic (PBPK) and pharmacokinetic (PK) models are routinely used in U.S. Environmental Protection Agency (EPA) toxicity assessments (U.S. EPA, 2024a, b) and represent the preferred approach for relating external exposure with internal measurements. The use of a PK model allows for RfD derivation based on concentrations that are measured within the body (e.g., in blood or urine) (U.S. EPA, 2014b, 2006).

For the future fluoride assessment, the EPA is considering existing PK models and approaches identified through its systematic review of the literature or public comment. Each PK or PBPK model is defined by a specific set of model equations and a number of model parameters which must be chosen appropriately to match existing physiological, pharmacokinetic, or in vivo data. In order to be used with a reasonable level of confidence in human health risk assessment, the model must first be evaluated for quality, to assure that:

1. It properly represents the underlying biology, given the assumptions stated or implied in the scientific reference(s) describing the model (the model equations are correct);
2. The model parameters taken directly from the scientific literature have been transcribed accurately and appropriately applied; and
3. The model otherwise fits or matches the appropriate ADME data available for model evaluation with an acceptable degree of precision, or sound explanations exist for data that the model does not fit such that the discrepancy can be reasonably ignored.

This approach stresses clarity in the documentation of model purpose, structure, and biological characterization; validation of mathematical descriptions, parameter values, and computer implementation; and evaluation of each plausible dose metric. Such transparency and documentation are important for compliance with the agency's information quality guidelines (U.S. EPA, 2002a) and are in line with previous PBPK model documentation publications (McLanahan et al., 2012; IPCS, 2010; Chiu et al., 2007; Clark et al., 2004).

The overall model evaluation criteria are split into two categories: scientific criteria and technical criteria. The scientific criteria focus on whether or not the biology, chemistry, and other information available for the chemical are appropriately represented by the model structure and equations. The scientific criteria can be judged based on the model documentation (publication or report) and do not require evaluation of the computer code. The technical criteria are assessed according to the computer code's precision and reliability in reflecting the scientific criteria and are subdivided into preliminary and in-depth categories. The preliminary technical criteria include availability of the computer code and apparent completeness of parameter listing and documentation. The in-depth technical criteria focus on the accurate implementation of the conceptual model in the model code and scripts, use of correct or biologically consistent parameters in the model, and reproducibility of model results reported in journal publications and other documents. The EPA is evaluating previously published fluoride pharmacokinetic models using the following scientific and technical criteria.

Scientific criteria for PBPK/PK models:

- Biological basis for the model is accurate
  - Model equations are consistent with biochemical understanding and biological plausibility
  - Consistent with mechanisms that significantly impact dosimetry
  - Describes critical behavior, such as nonlinear kinetics in a relevant dose range
  - Predicts dose metrics expected to be relevant and to be better correlated with toxicity or risk than applied doses
  - Applicable for relevant route(s) of exposure
- Model should describe existing PK data reasonably well
  - Shape: matches curvature or nonlinearity, inflection points, peak concentration time, etc.
  - Quantitative value: model predictions preferably within a factor of 2–3 of the data
- Validity of chemical-specific hypotheses:
  - Standard PBPK model compartments incorporate a limited number of hypotheses regarding ADME processes that have been tested and shown consistent with multiple data sets, for multiple chemicals, and therefore do not require in-depth consideration.
  - However, hypotheses specific to a particular chemical or chemical class, which are not supported by PBPK model agreement with data for other chemicals, should be evaluated more carefully, in particular when a hypothesis leads to prediction of much lower risk in humans than experimental animals.
    - For example, if it is hypothesized that a specific metabolic pathway operates in an experimental animal species (in a target tissue), making that species (tissue) particularly sensitive, then one should determine if there are ADME data for that metabolite (in the target tissue) in both sensitive and non-sensitive animal species demonstrating dosimetric differences commensurate with sensitivity, and dosimetric data in humans (or human tissues) demonstrating a lack of production.
    - Another example is the hypothesis that reactive metabolites formed in the liver will not have an impact on other tissues. But a moderately reactive metabolite with a half-life of minutes is sufficiently stable to be transported between tissues or cell types within a tissue, even if it is too reactive to measure in tissue samples from in vivo PK studies, so this hypothesis needs careful evaluation.
  - PBPK models which incorporate alternate hypotheses (e.g., some systemic distribution for a metabolite vs. none) may be equally consistent with the ADME data, but lead to very different risk predictions, and the resulting range of uncertainty should be considered.

Preliminary technical criteria for PBPK/PK models (evaluate if scientific criteria are met):

- Well-documented, available model code
- Parameters are clearly identified, including origin/derivation
- Parameters do not vary unpredictably with dose (e.g., any dose-dependence in absorption constants is predictable across the dose ranges relevant for animal and human modeling)
- Sensitivity and uncertainty analysis has been conducted for relevant exposure levels
    - If a sensitivity analysis was not conducted, then one may be performed as part of the evaluation
    - A sound explanation should be provided when sensitivity of the dose metric to model parameters differs from what is reasonably expected based on experience

In-depth technical criteria for PBPK/PK models (evaluate if the scientific and preliminary technical criteria are met):

- Model equations and parameters specified in computer code match those published or implied in the peer-reviewed manuscript or report
- For probabilistic human models, evaluate parameter distributions in the model vs. full human variability.
- Published figures and tables of model simulations are reproducible to within 10%
- If errors in the model implementation (equations or parameters) are found and corrected, and the correction or change alters the evaluated model predictions (plots or tables showing model agreement with data) by less than 10%, the error is considered small enough to not invalidate the model or any other parameter value, even if model predictions outside the range of the data change by more than 10%.
- Since model quality is judged by comparing model predictions to data, the impact of an error on model quality is evaluated only by determining the impact in the range of the data. The error is considered de minimis, hence acceptable, if the impact in the range of the data is less than 10%.
- An impact greater than 10% outside the range of any data may indicate uncertainty in model extrapolation to that range but does not alter the evaluation of its technical quality.
- If scientifically justified, a new version of the model equation or parameter may be documented and used in place of a published version (even if errors/corrections in the original version do not result in changes greater than 10%)
- For corrections resulting in changes greater than 10% in the range of the data, or significant changes in model structure (vs. only revising parameters), the revised model should be evaluated as a new model version; key conclusions may be unchanged, but the quality cannot be judged based on results of the previous version.