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

Proposed NTP Evaluation on Fluoride Exposure and Potential for Developmental Neurobehavioral Effects

November 19, 2015

Office of Health Assessment and Translation (OHAT)

Division of the National Toxicology Program

National Institute of Environmental Health Sciences

TABLE OF CONTENTS

PROPOSED NTP EVALUATION ON FLUORIDE EXPOSURE AND POTENTIAL DEVELOPMENTAL NEUROBEHAVIORAL EFFECTS

Project Leader:
Kristina Thayer, Office of Health Assessment and Translation (OHAT), DNTP

Summary: The National Toxicology Program (NTP) proposes to conduct an evaluation of the published literature to determine whether exposure to fluoride is associated with effects on neurodevelopment, specifically learning, memory, and cognition. This evaluation will use systematic review methods and include an examination of data from human (epidemiological), experimental animal, and mechanistic studies. Previous evaluations have found support for an association between fluoride exposure and impaired cognition; however, many of the studies included exposure to high levels of fluoride. Most of the human evidence was from fluoride-endemic regions having high background levels of fluoride, and the animal studies typically included exposure during development to relatively high concentrations of fluoride (>10 mg/L) in drinking water. Thus, the existing literature is limited in its ability to evaluate potential neurocognitive effects of fluoride in people associated with the current U.S. Public Health Service drinking water guidance (0.7 mg/L). In order to facilitate this literature-based evaluation, NTP is planning laboratory studies in experimental animals to address identified research needs and provide data useful for understanding effects of fluoride at water concentrations relevant to current human exposures. The findings from these studies will be included in the literature-based NTP evaluation of fluoride exposure and neurodevelopment.

BACKGROUND

EXPOSURE

Sources of fluoride exposure include drinking water, foods, beverages, dental products (toothpaste, mouth rinses), supplements, industrial emissions, pharmaceuticals, and pesticides (e.g., cryolite and sulfuryl fluoride). Soil ingestion is another source of exposure in young children (US Environmental Protection Agency 2010b).

In 2010, the US Environmental Protection Agency conducted a relative source contribution analysis and concluded that drinking water, beverages, food, and toothpaste are the major contributors to fluoride exposure (Table 1). The relative source contribution from drinking water intake was 40 to 60% after the age of 1 year and 70% in children less than 1 year old.

Table 1. Representative Values for Fluoride Intakes Used in Calculation of the Relative Source Contribution from Drinking Water

Age group (years)	DWI[a] (mg/day)	BI (mg/day)	FI (mg/day)	TI (mg/day)	SuF (mg/day)	SI (mg/day)	Total (mg/day)	RSC (%)
0.5 - <1	0.84	--	0.25[b]	0.07	0.03	0.02	1.2	70
1 - <4	0.63	0.36	0.16	0.34	0.05	0.04	1.58	40
4 - <7	0.82	0.54	0.35	0.22	0.06	0.04	2.03	40
7 - <11	0.86	0.60	0.41	0.18	0.07	0.04	2.16	40
11-14	1.23	0.38	0.47	0.20	0.09	0.04	2.41	51
>14	1.74[b]	0.59	0.38	0.10[c]	0.08	0.02	2.91	60

From Table 7-2 (US Environmental Protection Agency 2010b)
[a] Consumers only; 90th percentile intake except for >14 years. The >14 year value is based on the Office of Water policy of 2 L/day. [b] Includes foods, fluoride in powdered formula, and fruit juices; no allocation for other beverages. [c] Assumed. 50% of the 11-14 year old age group. DWI = Drinking Water Intake; BI = Beverage Intake; FI = Food Intake (Solid Foods); TI = Toothpaste Intake; SuF = Sulfuryl Fluoride Intake; SI = Soil Intake; RSC = Relative Source Contribution.

EPA has proposed a reference dose (RfD) of 0.08 mg/kg/day for protection against pitting of the tooth enamel (severe dental fluorosis) and this value is also considered protective against fractures and skeletal effects in adults (US Environmental Protection Agency 2010a). The RfD is the estimate of the daily exposure that is likely to be without harmful effect during a lifetime. A RfD of 0.08 is equivalent to a daily dose of 5.6 mg for a 70 kg person or 1.6 mg for a 20 kg child.

USE OF FLUORIDE TO PREVENT TOOTH DECAY

Fluoride from community water fluoridation, mouth rinses, gels and toothpastes is intended to prevent dental caries primarily through topical remineralization of tooth surfaces. Community water fluoridation and fluoride toothpaste are the most common sources of non-dietary fluoride in the United States (U.S. DHHS Federal Panel on Community Water Fluoridation 2015). Because fluorine is the 13th most abundant element in the earth's crust, fluoride is also naturally occurring in water, and is present even in non-fluoridated water systems.

Although other fluoride-containing products and sources are available (e.g., mouth rinses, dietary supplements, professionally applied fluoride compounds), community water fluoridation has been identified as the most cost-effective method of delivering fluoride to all members of the community regardless of age, educational attainment, or income level. Consuming fluoridated water and beverages, and foods prepared or processed with fluoridated water throughout the day maintains a low concentration of fluoride in saliva and plaque that enhances remineralization. Community water fluoridation to minimize the occurrence and severity of tooth decay began in 1945 and by 2012 had reached 67% of the U.S. population. About 25 countries practice community water fluoridation (Iheozor-Ejiofor et al. 2015) and many more countries provide fluoride through other means such as salt. In 2012, an estimated 200 million people in the U.S. were served by 12,341 community water systems that added fluoride to water or purchased water with added fluoride from other systems (U.S. DHHS Federal Panel on Community Water Fluoridation 2015).

The U.S. Public Health Service (PHS) first recommended communities add fluoride to drinking water in 1962. PHS guidance is advisory, not regulatory, which means that while PHS recommends community water fluoridation as an effective public health intervention, the decision to fluoridate water systems is made by state and local governments. For community water systems that add fluoride, PHS now recommends a fluoride concentration of 0.7 milligrams/liter (mg/L)[1] based on the optimal concentration of fluoride in drinking water. This recommended level provides the best balance of protection from dental caries while limiting the risk of dental fluorosis, a condition marked by changes in the appearance of tooth enamel most commonly appearing as lacy white markings (U.S. Department of Health and Human Services Federal Panel on Community Water Fluoridation 2015). Dental fluorosis may result when children regularly consume fluoride from birth through 8 years of age -- the time that their permanent teeth (with the exception of the third molars) are developing.

Under the Safe Drinking Water Act, the U.S. Environmental Protection Agency (EPA) sets standards for drinking water quality. Currently, the enforceable standard is set at 4.0 mg/L to protect consumers from exposure to drinking water sources with naturally high occurrence of fluoride against severe skeletal fluorosis (i.e., a condition caused by excessive fluoride intake for a long period of time that in advanced stages can cause pain and/or crippling damage to bones and joints). EPA also has a secondary drinking water standard of 2.0 mg/L to protect against moderate to severe dental fluorosis, which is not enforceable but requires systems to notify the public. The EPA is in the process of reviewing the current drinking water standards for fluoride (US EPA 2013).

CONCERNS FOR POTENTIAL FLUORIDE TOXICITY

The NTP received a nomination in June 2015 from the public to conduct an analysis of fluoride developmental neurobehavioral toxicity in June 2015. Concerns for possible adverse health effects of fluoride were also raised in public comments received on the Proposed Recommendation for Fluoride Concentration in Drinking Water for the Prevention of Dental Caries published in 2011 (U.S. DHHS Federal Panel on Community Water Fluoridation 2015). Commonly cited health concerns raised in the public comments included bone fractures and skeletal fluorosis, IQ and other neurological effects, and

[1] For many years most community water fluoridated systems used fluoride concentrations ranging from 0.8 to 1.2 mg/L (U.S. Department of Health and Human Services Federal Panel on Community Water Fluoridation 2015)

cancer and endocrine disruption. Both cancer and endocrine disruption have also been nominated to the NTP for evaluation.

Effects on neurological function, endocrine (thyroid, parathyroid, pineal, glucose metabolism), and carcinogenicity were assessed in the 2006 National Research Council (NRC) report "Fluoride in Drinking Water: A Scientific Review of EPA's Standards (National Research Council 2006), which considered adverse effects of water fluoride focusing on concentrations of 2–4 mg/L,[2] a level higher than that currently recommended for community water fluoridation (0.7 mg/L). At levels below 4.0 mg/L, NRC found no evidence substantial enough to support negative health effects other than severe dental fluorosis. The conclusions from the 2006 NRC review were accepted as the summary of potential hazard in the 2015 U.S. Department of Health and Human Services (DHHS) Federal Panel on Community Water Fluoridation report. As part of its review, the NRC report noted a number of challenges to evaluating the literature for each of these topic areas, including deficiencies in reporting quality; consideration of all sources of exposure to fluoride, confounding, and use of similar comparison populations in the epidemiology studies; clinical significance of endocrine effects, and the relationship between histological, biochemical, and molecular changes and alterations in behavior or disease status. The NRC report also presented a series of research recommendations.

The main conclusions with respect to neurotoxicity and neurobehavioral effects in the 2006 NRC report were:

- "Animal and human studies of fluoride have been published reporting adverse cognitive and behavioral effects. A few epidemiologic studies of Chinese populations have reported IQ deficits in children exposed to fluoride at 2.5 to 4 mg/L in drinking water. Although the studies lacked sufficient detail for the committee to fully assess their quality and relevance to U.S. populations, the consistency of the results appears significant enough to warrant additional research on the effects of fluoride on intelligence." [p. 8]

- "A few animal studies have reported alterations in the behavior of rodents after treatment with fluoride, but the committee did not find the changes to be substantial in magnitude. More compelling were studies on molecular, cellular, and anatomical changes in the nervous system found after fluoride exposure, suggesting that functional changes could occur. These changes might be subtle or seen only under certain physiological or environmental conditions. More research is needed to clarify the effect of fluoride on brain chemistry and function." [p. 8]

Since release of the 2006 NRC report, approximately 10 epidemiological studies of children's IQ and over 40 experimental animal studies related to developmental neurological effects of fluoride have been published. Updated analyses by the European Commission's Scientific Committee on Health and Environmental Risks (SCHER 2011) and the Royal Society of New Zealand/Office of the Prime Minister's Chief Science Advisor (2014) did not find support from human studies for effects on children

[2] EPA's maximum-contaminant-level goal (MCLG) for fluoride is 4mg/L and secondary maximum contaminant level (SMCL) is 2 mg/L. The goal of the MCLG is to establish an exposure guideline to prevent adverse health effects in the general population, and the goal of the SMCL is to reduce the occurrence of adverse cosmetic consequences from exposure to fluoride. Both the MCLG and the SMCL are non-enforceable guidelines (NRC 2006).

neurodevelopment at levels of fluoride in drinking water. However, neither of these reviews used systematic review methodology and neither was comprehensive in identifying and describing relevant animal studies. A 2015 systematic analysis of the human literature conducted for the Republic of Ireland's Department of Health (Sutton et al. 2015) concluded that there was no evidence of an association with lowered IQ in studies of community water fluoridation areas based primarily on an analysis of a prospective cohort study in New Zealand (Broadbent et al. 2015). For fluoride-endemic areas, there was a strong suggestion that high levels of naturally occurring fluoride in water (> 1.5 ppm) may be associated with negative health effects, including lowering of IQ. In general, these studies were considered of low quality because they did not fully account for other factors that could also cause a lowering of IQ e.g., nutritional status, socioeconomic status, iodine deficiency, other chemicals in the ground water (arsenic or lead). The conclusions of Sutton et al. (2015) are consistent with findings of a 2012 meta-analysis of 27 epidemiology studies that supported the possibility of an adverse effect of "high" fluoride exposure[3] on children's neurodevelopment, specifically for lowered IQ; although the 2012 meta-analysis also identified study quality limitations, mostly related to reporting quality, that limited the strength of conclusions that could be reached (Choi et al. 2012).

The NTP has recently completed a systematic review of fluoride and neurobehavioral outcomes in animal studies that included consideration of adult and developmental exposure and a broad range of behavioral outcomes, including learning and memory, motor and sensory function, depression and motor endurance, anxiety and motor activity. This report is currently undergoing peer-review and expected to be finalized early in 2016. A total of 61 studies were considered relevant (Appendix A), and 44 of these addressed learning and memory. For evidence synthesis, 14 of the learning and memory studies were excluded based on serious concern for risk of bias (internal validity), leaving a total of 30 studies considered in an analysis of learning and memory in rats and mice. Draft conclusions found evidence of potential detrimental effects on learning and memory, but confidence in the conclusions was limited due to study design and reporting issues and there was also concern for potential confounding of the learning and memory assessments by deficits in motor function or fear responses. Most of the studies reporting effects treated animals with doses >10 ppm. Few studies tested dose levels of less than 5 ppm (Zhang et al. 1999; Xu and Shen 2001; Gao et al. 2008; Gao et al. 2009; Liu et al. 2009; Liu et al. 2010; Liu et al. 2011; Zhu et al. 2012; Liu et al. 2014) and none of these assessed the impacts of exposure during development (Appendix A). Further, levels of fluoride in vehicle controls in the lower dose studies (<10 ppm) ranged from 0.15 to 0.7 ppm, at or only slightly lower than the current PHS guidance (Chioca et al. 2008; Gao et al. 2008; Gui et al. 2010; Jiang et al. 2014; Wei et al. 2014). For these reasons, the animal literature on learning and memory following developmental exposure is not considered sufficient to assess effects at dose levels relevant to current water fluoridation practices in the US. The draft report concludes that additional studies are required to have higher confidence in the specificity of the responses as learning or memory impairments and in quantitative measures such as identification of No Observed Effect Level (NOEL) or Lowest Observed Effect Level doses, or parameters for benchmark dose analysis. The NTP is currently pursuing experimental studies in rats to address key data gaps, starting with pilot studies that address limitations of the current literature with respect to study design (e.g., randomization, blinding, control for litter effects), and assessment of motor and sensory function to assess the degree to which impairment of movement may impact performance in

[3] "High" was defined based on drinking water concentration, evidence of fluorosis, exposure related to coal-burning activities, and urine levels.

learning and memory tests. If justified, follow-up studies would address potential developmental effects using lower dose levels more applicable to human intakes.

Given the number of studies published since the 2006 NRC and 2011 SCHER evaluations, there appears to be sufficient rationale to justify conducting an evaluation that integrates evidence from epidemiological, experimental animal, and mechanistic[4] data to reach an NTP hazard identification conclusion with respect to developmental neurobehavioral toxicity. However, an analysis of the existing literature would likely be limited in its ability to reach conclusions about potential cognitive effects in people associated with the current drinking water guidance (0.7 mg/L). For this reason, the timing of the analysis will be structured to include the results of the experimental animal studies currently being initiated by the NTP. This should enable a more complete interpretation of the animal data with respect to understanding potential neurocognitive effects at water concentrations relevant to current human exposure levels.

With respect to evaluations of cancer and non-thyroid endocrine outcomes, separate analyses are proposed to determine the amount of evidence available and merit of pursuing systematic reviews given factors such as the extent of new research published since previous evaluations, and whether these new reports address or correct the deficiencies noted in the literature (National Research Council 2006; California Office of Environmental Health Hazard Assessment (OEHHA) 2011; Scientific Committee on Health and Environmental Risks (SCHER) 2011).

OBJECTIVE AND SPECIFIC AIMS

The overall objective of this evaluation is to undertake an integrated analysis of evidence from human, animal, and mechanistic studies to develop hazard identification conclusions about whether fluoride is a developmental neurobehavioral toxicant. The evaluation will be implemented by developing a protocol based on guidance in the OHAT Handbook for Systematic Review and Evidence Integration (NTP 2015).

Steps in the process and specific aims:

- Identify literature reporting the effects of exposure to fluoride and developmental neurological outcomes in humans, non-human mammals, or in applicable *in vitro* and *in silico*[5] model systems.
- Extract data on health outcomes from relevant studies.
- Assess the internal validity ("risk of bias") of individual human and non-human mammalian studies.

[4] Mechanistic data come from a wide variety of studies and are generally not intended to identify a disease phenotype. This source of experimental data includes *in vitro* and *in vivo* laboratory studies directed at identifying the cellular, biochemical, and molecular mechanisms that are related to chemicals that produces particular adverse effects. These studies increasingly take advantage of new "-omics" tools, such as proteomics and metabolomics, to identify early biomarkers of effect. Another broad class of mechanistic data relates to the toxicokinetics of a chemical (NRC 2014).

[5] *In silico* refers to computer-based models

- Summarize the extent and types of evidence available.

The following specific aims will depend on the extent and nature of the available evidence (i.e., number and similarity of studies):

- Synthesize the evidence, including performance of quantitative meta-analyses if appropriate, and evaluate sources of heterogeneity.
- Rate confidence in the body of evidence for neurological effects for human and non-human mammalian studies separately according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Very Low/No Evidence Available.
- Translate confidence ratings into level of evidence of health effects for human and non-human mammalian studies separately according to one of four statements: (1) High, (2) Moderate, (3) Low, or (4) Inadequate.
- Combine the level of evidence ratings for human and non-human mammalian data and consider the degree of support from mechanistic data to reach one of five possible hazard identification conclusions: (1) Known, (2) Presumed, (3) Suspected, (4) Not classifiable, or (5) Not identified to be a hazard to humans.
- Describe findings in the context of human exposure levels, describe limitations of the analysis, and identify data gaps and key research needs.

DRAFT PECO STATEMENT

A PECO (**P**opulation, **E**xposure, **C**omparators and **O**utcomes) statement (Table 2) is used as an aid to focus the research question(s), search terms, and inclusion/exclusion criteria in a systematic review (Higgins and Green 2011). The draft PECO statement was based on a series of problem formulation steps that included: (1) review of the nomination, (2) discussion with staff at Federal agencies and the nominator; (3) consultation with an evaluation design team[6] with expertise in neurotoxicology, epidemiology, toxicology, systematic review and evidence integration, and information science; and (4) consideration of information received from a public request for information in the Federal Register [80 FR 60692 (October 7, 2015) 60692 -60693].

[6] The evaluation team is composed of NIEHS/NTP staff, staff from other US Federal agencies, and contractor staff who are involved in the entire systematic review process. As needed, OHAT will also engage non-federal technical advisors, who are screened for potential conflicts of interest. Contractor staff members are also screened for potential conflicts of interest.

Table 2. Draft PECO (Population, Exposure, Comparators and Outcomes) statement	
PECO Element	**Evidence**
Population	Humans, non-human mammalian animal species (whole organism, *ex vivo*), and *in vitro or in silico* model systems.
Exposure	**Forms of fluoride (CASRN):** Sodium fluoride (7681-49-4, the most common form used in toxicology studies), soluble fluorine (7782-41-4), fluorosilicic acid (16961-83-4), or sodium fluorosilicate (16893-85-9). **Humans and non-human mammalian animal species:** Fluoride exposure or treatment that includes a developmental life-stage, i.e., during fetal life, infancy, childhood (i.e., ≤18 years in humans; up to post-natal day 30 in rodent species). There are no restrictions based on dose level (in order to help assess shape of dose response). **In vitro/in silico models:** Fluoride treatment with no restrictions on life-stage of model system.
Comparators	**Humans:** A comparison group exposed to no or lower levels compared to more highly exposed participants. **Non-human mammalian animal species**: Experimental study that includes a vehicle-only control treatment. **In vitro/in silico models:** Experimental tissue, cell, or cell component study that includes a vehicle-only control treatment for *in vitro* studies; comparison group not required for in silico models.
Outcomes	**Humans and non-human mammalian animal species**: *Primary outcomes:* Neurobehavioral outcomes related to cognition *Secondary (mechanistic) outcomes:* Brain-related cellular, morphometric or histological endpoints; thyroid hormone-related measures; toxicokinetic data. **In vitro/in silico models:** *Secondary (mechanistic) outcomes:* Brain-related endpoints in studies of neuronal cells, neurotransmitters, and/or receptors.

CONSIDERATIONS FOR PROTOCOL DEVELOPMENT AND ANALYSIS

After considering public comments on the draft concept document, a detailed protocol will be developed following guidance outlined in the OHAT Handbook for Systematic Review and Evidence Integration (NTP 2015). The protocol will be posted on the OHAT website. Any revisions during the course of the evaluation will be noted. The following section is intended to highlight key issues that will be considered when developing the study protocol. The protocol and draft report will be developed by NTP staff, other members of the evaluation design team, and technical advisors (as needed) who have been screened for conflict of interest.

LITERATURE SEARCH STRATEGIES

Literature search strategies will be developed to identify published evidence on the effects of fluoride on neurological outcomes by using index terms and text words based on key elements of the research question. Six electronic databases[7] will be searched:

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

No publication date or language restrictions will be applied.

The reference lists of included studies, reviews related to neurological effects of fluoride, and the other compilations of studies related to fluoride (e.g., received through public comments, Fluoride Action Network database) will be searched for additional relevant publications. The list of included (and excluded) studies will be posted on the OHAT website prior to release of a draft report as an additional strategy to identify potentially relevant studies that may have been missed during the literature search.

[7] The National Library of Medicine's Toxline database is not included in the search because recent changes have resulted in significant reductions in search functionality that limits running the search strings for this topic. The other databases proposed for searching are very likely to identify relevant published and peer-reviewed animal studies. In addition, three other databases were searched in a prior NTP report ("Systematic Review on the Neurobehavioral Effects of Fluoride in Animal Studies," currently under internal review) and no relevant records were identified, thus they will not be searched in the current project: European Chemicals Agency (ECHA) Registration dossiers ("REACH"); Organization for Economic Co-operation and Development (OECD) Existing Chemicals Screening Information Data Sets (SIDS); USEPA HPV Challenge Program Robust Summaries and Test Plans.

SELECTION CRITERIA FOR THE EVIDENCE

In order to be eligible for inclusion, studies will need to comply with the criteria specified by the PECO statement (Table 2). Studies that do not meet the PECO criteria will be excluded. In addition, the following exclusion criteria will be applied:

- Studies do not contain original data, such as reviews, editorials, or commentaries.
    - Reference list of reviews were reviewed to identify potentially relevant articles.
- Studies not containing sufficient detail to undergo peer-review (e.g., conference abstracts, unpublished data described in technical reports, databases, working papers from research groups or committees, and white papers).
- Unpublished or non-peer-reviewed data that cannot be made publically available (see below for guidance).

Unpublished or non-peer-reviewed data

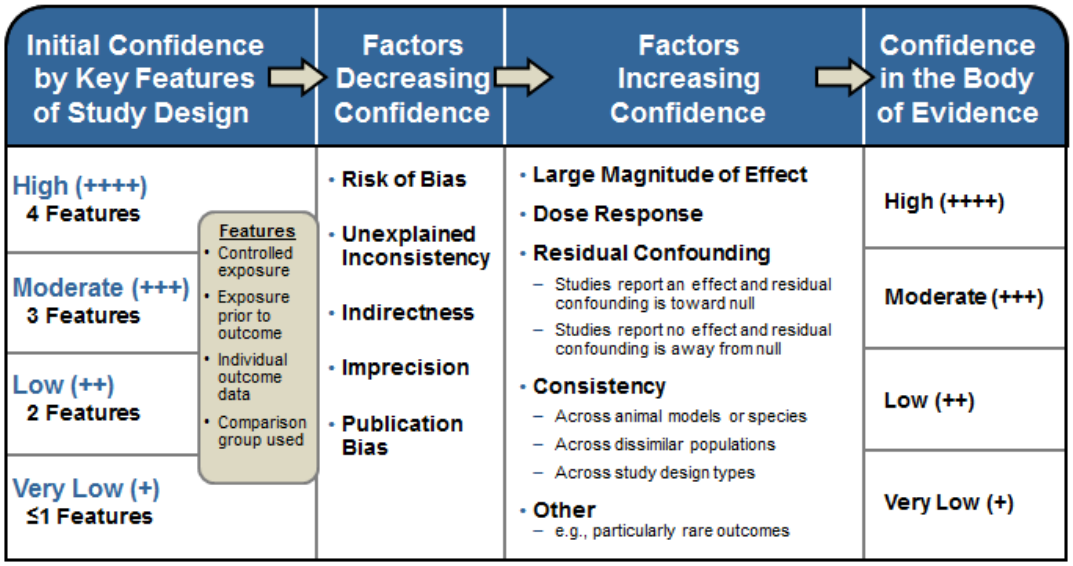
NTP only includes publicly accessible information in its evaluations. This information is typically based on studies published in peer-reviewed journals. However, NTP can consider unpublished data or data presented in the grey literature (e.g., theses/dissertations, technical reports, white papers) that has not undergone peer-review provided the owners of the data are willing to have the study details and results made publicly accessible. Peer-review of this data would be accomplished using standard procedures in the OHAT handbooks to evaluate the quality of the information with the option to utilize topic specific technical advisors as needed. Study sponsors and researchers are invited to submit unpublished data during the course of an evaluation, although the ability to use the information depends on the timing of submission relative to release of a draft monograph. Unpublished data from personal author communication can supplement a peer-reviewed study, as long as the information is made publicly available.

ASSESSMENT OF CONFIDENCE IN THE BODY OF EVIDENCE

In more complete description of the process and guidance used to implement the analysis is outlined in the OHAT Handbook for Systematic Review and Evidence Integration (NTP 2015). In brief, the quality of evidence for each outcome will be graded using the GRADE system for rating the confidence in the body of evidence (Guyatt et al 2011) as adapted by OHAT for observational human studies and animal studies (NTP 2015). Under the GRADE system, the overall confidence in the body of evidence for an outcome is categorized as high, moderate, low or very low. An initial confidence rating for the body of evidence (for a specific outcome) is determined by the ability of the study design to ensure that exposure preceded and was associated with the outcome (Figure 1, column 1). This ability is reflected in the presence or absence of four key study design features used to delineate the studies for initial confidence ratings: (1) the exposure to the substance is experimentally controlled, (2) the exposure assessment demonstrates that exposures occurred prior to the development of the outcome (or concurrent with aggravation/amplification of an existing condition), (3) the outcome is assessed on the individual level (i.e., not through population aggregate data), and (4) an appropriate comparison group is included in the study. The first key feature, "controlled exposure," reflects the ability of experimental studies in humans and animals to largely eliminate confounding by randomizing allocation of exposure. Therefore, these

studies usually have all four features and receive an initial rating of "High Confidence." Observational studies do not have controlled exposure and are differentiated by the presence or absence of the three remaining study design features. For example, prospective cohort studies usually have all three remaining features and receive an initial rating of "Moderate Confidence". Next, a series of adjustments ("downgrades" or "upgrades") may be made to the initial ranking based on the characteristics of the studies constituting the body of evidence after considering factors such as risk of bias across studies, unexplained inconsistency, indirectness, imprecision, publication bias, magnitude of the effect, dose response, and consistency across different model systems and study designs (Figure 1). Studies conducted in mammalian model systems are assumed relevant for humans (i.e., not downgraded for indirectness) unless compelling evidence to the contrary exist.

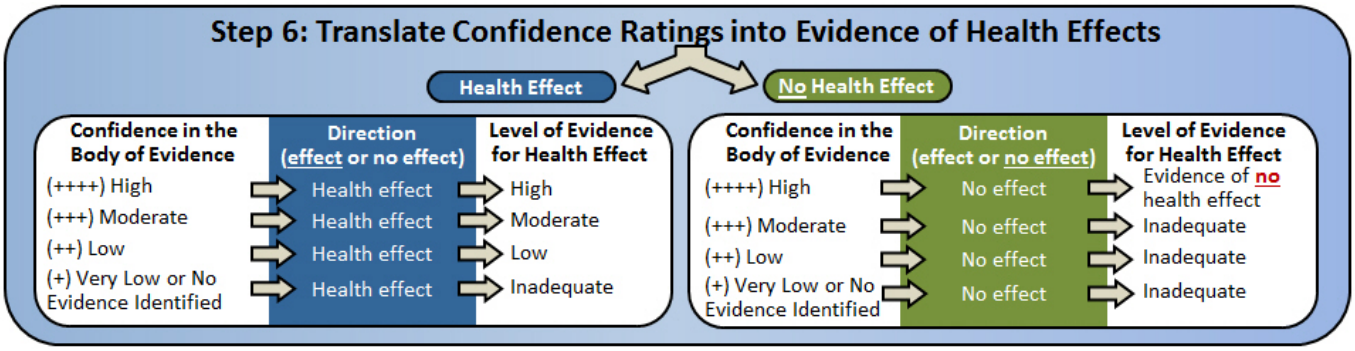
Figure 1. Assessing Confidence in the Body of Evidence



PREPARATION OF LEVEL OF EVIDENCE CONCLUSIONS

The confidence in the body of evidence conclusions from Figure 1 are translated into draft statements of health effects for human and animal data, seperately, according to one of four statements: 1. High, 2. Moderate, 3. Low, or 4. Inadequate (Figure 2, labeled as Step 6 in OHAT's process for systematic review and evidence integration). The descriptor "evidence of no health effect" is used to indicate confidence that the substance is not associated with a health effect. Because of the inherent difficulty in proving a negative, the conclusion "evidence of no health effect " is only reached when there is high confidence in the body of evidence.

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



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

INTEGRATE EVIDENCE TO DEVELOP HAZARD IDENTIFICATION CONCLUSIONS

For determining the appropriate hazard identification category, the evidence streams for human studies and animal studies, which have remained separate through the previous steps, are integrated along with other relevant data, such as supporting evidence from *in vitro* studies.

Integration of human and animal evidence

Hazard identification conclusions are initially reached by integrating the highest level-of-evidence conclusion for a health effect(s) from the human and the animal evidence streams. On an outcome basis, this approach applies to whether the data support a health effect conclusion or provide evidence of no health effect. The five hazard identification conclusion categories are as follows:

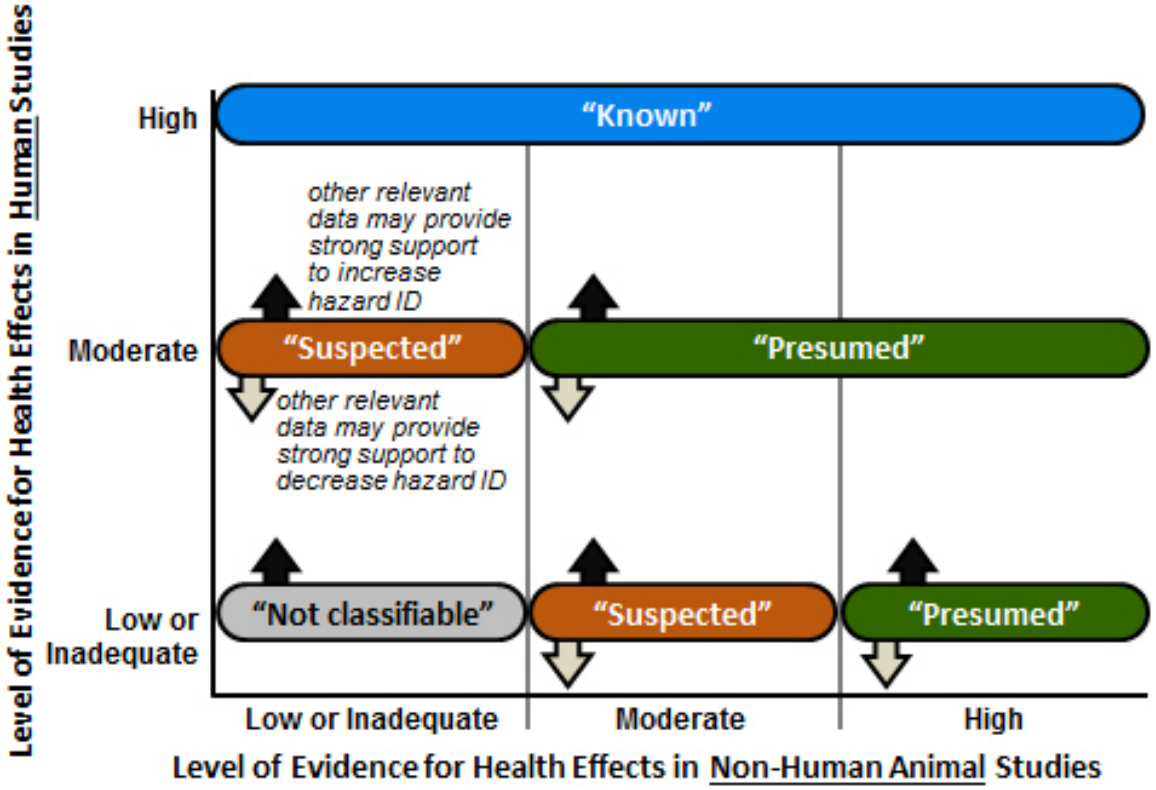
- Known to be a hazard to humans
- Presumed to be a hazard to humans
- Suspected to be a hazard to humans
- Not classifiable as a hazard to humans
- Not identified as a hazard to humans

When the data support a health effect, the level-of-evidence conclusion for human data from Step 6 is considered together with the level of evidence for non-human animal data to reach one of four hazard identification conclusions (Figure 3, labeled as Step 7 in OHAT's process for systematic review and evidence integration). If one evidence stream (either human or animal) is characterized as "Inadequate Evidence," then conclusions are based on the remaining evidence stream alone (which is equivalent to treating the missing evidence stream as "Low" in Step 7).

Consideration of mechanistic data

The NTP does not require mechanistic or mode-of-action data in order to reach hazard identification conclusions, although when available, this and other relevant supporting types of evidence may be used to raise (or lower) the category of the hazard identification conclusion. If mechanistic data provide strong support for biological plausibility of the relationship between exposure and the health effect, the hazard identification conclusion may be upgraded (indicated by black "up" arrows in the Step 7 graphic in Figure 3) from the one initially derived by considering the human and non-human animal evidence together.

Figure 3. Hazard Identification Scheme



IDENTIFICATION AND EVALUATION OF MOST RELEVANT MECHANISTIC STUDIES

Human and experimental animal data will be interpreted in conjunction with evidence from mechanistic data to evaluate the biological plausibility of any associations between fluoride and developmental

neurological effects. Relevant mechanistic evidence will be identified and evaluated using an iterative approach adapted from the US EPA Handbook of Procedures for Systematic Review In support of Integrated Risk Information System (IRIS) Toxicological Reviews (presented at November 17-18, 2015 National Academy of Sciences meeting "Unraveling Low Dose Toxicity: Case Studies of Systematic Review of Evidence").

- **Identification and categorization of mechanistic literature:** *In vitro* or *in silico* studies identified in the initial neurotoxicity-focused literature search will be tagged to develop a "bin" for mechanistic studies. Full-text review of studies in humans and non-human mammalian animal species will be conducted to determine if they also contain mechanistic data. Studies in non-mammalian animal species (e.g., fish, *C. elegans*) will be considered supportive information to assess biological plausibility and categorized as mechanistic.
- **Identification of proposed mechanism of action (MOAs) or mechanistic hypotheses from published literature:** The evaluation team will review the bibliographic information gathered from the literature survey of human, animal and *in vitro* studies to identify emerging patterns of potential neurotoxicity. These patterns will inform hypothesized mechanistic events. Additional targeted literature search protocols may be conducted to identify other potentially relevant mechanisms if needed.
- **Prioritization of mechanistic studies for analysis:** Once neurological effects of interest are identified from the human and animal studies, the evaluation team will evaluate the mechanistic data to focus on the studies and outcomes that are most informative for those outcomes. The protocol will be updated to indicate which types of mechanistic studies are considered most relevant.
- **Evaluation:** After prioritization, the most relevant set(s) of experimental studies will be evaluated. For topics with large evidence base, reviews by others may be used. Studies should be grouped by assay and/or endpoint type to facilitate analysis of support for biological plausibility.

REFERENCES

[Anonymous]. 2014. Health effects of water fluoridation: A review of the scientific evidence. A report on behalf of the Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor. Available at http://www.royalsociety.org.nz/media/2014/08/Health-effects-of-water-fluoridation_Aug_2014_corrected_Jan_2015.pdf [accessed 9 November 2015].

Broadbent JM, Thomson WM, Ramrakha S, Moffitt TE, Zeng J, Foster Page LA, Poulton R. 2015. Community Water Fluoridation and Intelligence: Prospective Study in New Zealand. Am J Public Health 105(1):72-76 PMC - PMC4265943 PMCR- 4262016/4265901/4265901 4265900 4265900.

California Office of Environmental Health Hazard Assessment (OEHHA). 2011. California Office of Environmental Health Hazard Assessment. Meeting synopsis and slide presentations: carcinogen identification committee meeting held on October 12, 2011. Available from: URL: http://oehha.ca.gov/prop65/public_meetings/cic101211synop.html [accessed 17 September 2015].

Chioca LR, Raupp IM, Da Cunha C, Losso EM, Andreatini R. 2008. Subchronic fluoride intake induces impairment in habituation and active avoidance tasks in rats. Eur J Pharmacol 579(1-3):196-201.

Choi AL, Sun GF, Zhang Y, Grandjean P. 2012. Developmental Fluoride Neurotoxicity: A Systematic Review and Meta-Analysis. Environ Health Perspect 120(10):1362-1368.

Gao Q, Liu YJ, Guan ZZ. 2009. Decreased learning and memory ability in rats with fluorosis: Increased oxidative stress and reduced cholinesterase activity in the brain. Fluoride 42(4):277-285.

Gao Q, Liu YJ, Wu CX, Long YG, Guan ZZ. 2008. Effects of fluoride on learning and memory and cholinesterase activity in rat brains. Chinese Journal of Endemiology 27(2):128-130.

Gui CZ, Ran LY, Li JP, Guan ZZ. 2010. Changes of learning and memory ability and brain nicotinic receptors of rat offspring with coal burning fluorosis. Neurotoxicol Teratol 32(5):536-541.

Higgins J, Green S. 2011. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011). http://handbook.cochrane.org/ [accessed 3 February 2013].

Iheozor-Ejiofor Z, Worthington HV, Walsh T, O'Malley L, Clarkson JE, Macey R, Alam R, Tugwell P, Welch V, Glenny AM. 2015. Water fluoridation for the prevention of dental caries. The Cochrane database of systematic reviews 6:CD010856.

Jiang C, Zhang S, Liu H, Guan Z, Zeng Q, Zhang C, Lei R, Xia T, Wang Z, Yang L, Chen Y, Wu X, Zhang X, Cui Y, Yu L, Wang A. 2014. Low glucose utilization and neurodegenerative changes caused by sodium fluoride exposure in rat's developmental brain. NeuroMolecular Medicine 16(1):94-105.

Liu F, Ma J, Zhang H, Liu P, Liu YP, Xing B, Dang YH. 2014. Fluoride exposure during development affects both cognition and emotion in mice. Physiol Behav 124:1-7.

Liu YJ, Gao Q, Long YG, Yu YN, Guan ZZ. 2011. Influence of chronic fluorosis on expression of phospho-Elk-1 in rat brains. Chinese Journal of Endemiology 30(3):251-255.

Liu YJ, Gao Q, Wu CX, Guan ZZ. 2010. Alterations of nAChRs and ERK1/2 in the brains of rats with chronic fluorosis and their connections with the decreased capacity of learning and memory. Toxicol Lett 192(3):324-329.

Liu YJ, Gao Q, Wu CX, Long YG, Guan ZZ. 2009. Modified expression of extracellular signal-regulated protein kinase signal transduction in rat brains and changed capacity of learning and memory of rats with chronic fluorosis. Chinese Journal of Endemiology 28(1):32-35.

National Research Council. 2006. Committee on Fluoride in Drinking Water, Board on Environmental Studies and Toxicology. Fluoride in drinking water: a scientific review of EPA's standards. Washington: National Academies Press. Available at http://www.nap.edu/catalog/11571/fluoride-in-drinking-water-a-scientific-review-of-epas-standards [cited 2015 August 23]. .

NRC (National Research Council). 2014. Review of EPA's Integrated Risk Information System (IRIS) Process (http://www.nap.edu/catalog.php?record_id=18764) [accessed 1 January 2015].

NTP (National Toxicology Program). 2015. Handbook for Conducting a Literature-Based Health Assessment Using Office of Health Assessment and Translation (OHAT) Approach for Systematic Review and Evidence Integration. January 9, 2015 release. Available at http://ntp.niehs.nih.gov/go/38673.

Scientific Committee on Health and Environmental Risks (SCHER). 2011. European Commission Directorate-General for Health and Consumers, Scientific Committees. Critical review of any new evidence on the hazard profile, health effects, and human exposure to fluoride and the fluoridating agents of drinking water. Available at http://ec.europa.eu/health/scientific_committees/environmental_risks/docs/scher_o_139.pdf [accessed 17 September 2015].

Sutton M, Kiersey R, Farragher L, Long J. 2015. Health Effects of Water Fluoridation: An evidence review. Report conducted for Republic of Ireland's Department of Health. Availalbe at http://www.hrb.ie/uploads/tx_hrbpublications/Health_Effects_of_Water_Fluoridation.pdf [accessed 9 November, 2015].

U.S. Department of Health and Human Services Federal Panel on Community Water Fluoridation. 2015. PHS Recommendation for Fluoride Concentration in Drinking Water. Available at http://www.publichealthreports.org/documents/PHS_2015_Fluoride_Guidelines.pdf [accessed 17 September 2015]. Public Health Reports 130(July-August):318-331.

US Environmental Protection Agency. 2013. Basic information about fluoride in drinking water: Review of fluoride drinking water standard. Available at http://www2.epa.gov/dwsixyearreview/review-fluoride-drinking-water-regulation [accessed 16 November, 2015].

US Environmental Protection Agency. 2010a. Fluoride: Dose-Response Analysis For Non-cancer Effects . 820-R-10-019. Washington: EPA, Office of Water, Health and Ecological Criteria Division. Available at

http://water.epa.gov/action/advisories/drinking/upload/Fluoride_dose_response.pdf cited 2015 October 25]. .

US Environmental Protection Agency. 2010b. Fluoride: exposure and relative source contribution analysis. 820-R-10-015. Washington: EPA, Office of Water, Health and Ecological Criteria Division. Available at http://water.epa.gov/action/advisories/drinking/upload/fluoridereport.pdf [cited 2015 August 23].

Wei N, Dong Y, Wang Y, Guan Z. 2014. Effects of chronic fluorosis on neurobehavioral development in offspring of rats and antagonistic effect of Vitamin E. Chinese Journal of Endemiology 33(2):125-128.

Xu X, Shen X. 2001. Effect of fluorosis on mice learning and memory behaviors and brain SOD activity and MDA content [pdf title: The influence of Fluoresis on Learning-memory Behaviors of Mice And the Activities of SOD and the Content of MDA in the Brain]. China Public Health 17:8-10.

Zhang Z, Xu X, Shen X, Xu X. 1999. [Effect of fluoride exposure on synaptic structure of brain areas related to learning-memory in mice]. Wei Sheng Yan Jiu 28(4):210-212.

Zhu YL, Zheng YJ, LV XM, Ma Y, Zhang J. 2012. Effects of fluoride exposure on performance in water labyrinth and monoamine neurotransmitters of rats. Journal of Xinjiang Medical University 35.

APPENDIX A. STUDY FLOW AND OVERVIEW OF INCLUDED STUDIES IN DRAFT NTP SYSTEMATIC REVIEW ON THE NEUROBEHAVIORAL TOXICITY OF FLUORIDE IN ANIMAL STUDIES

Figure S1. Study Flow

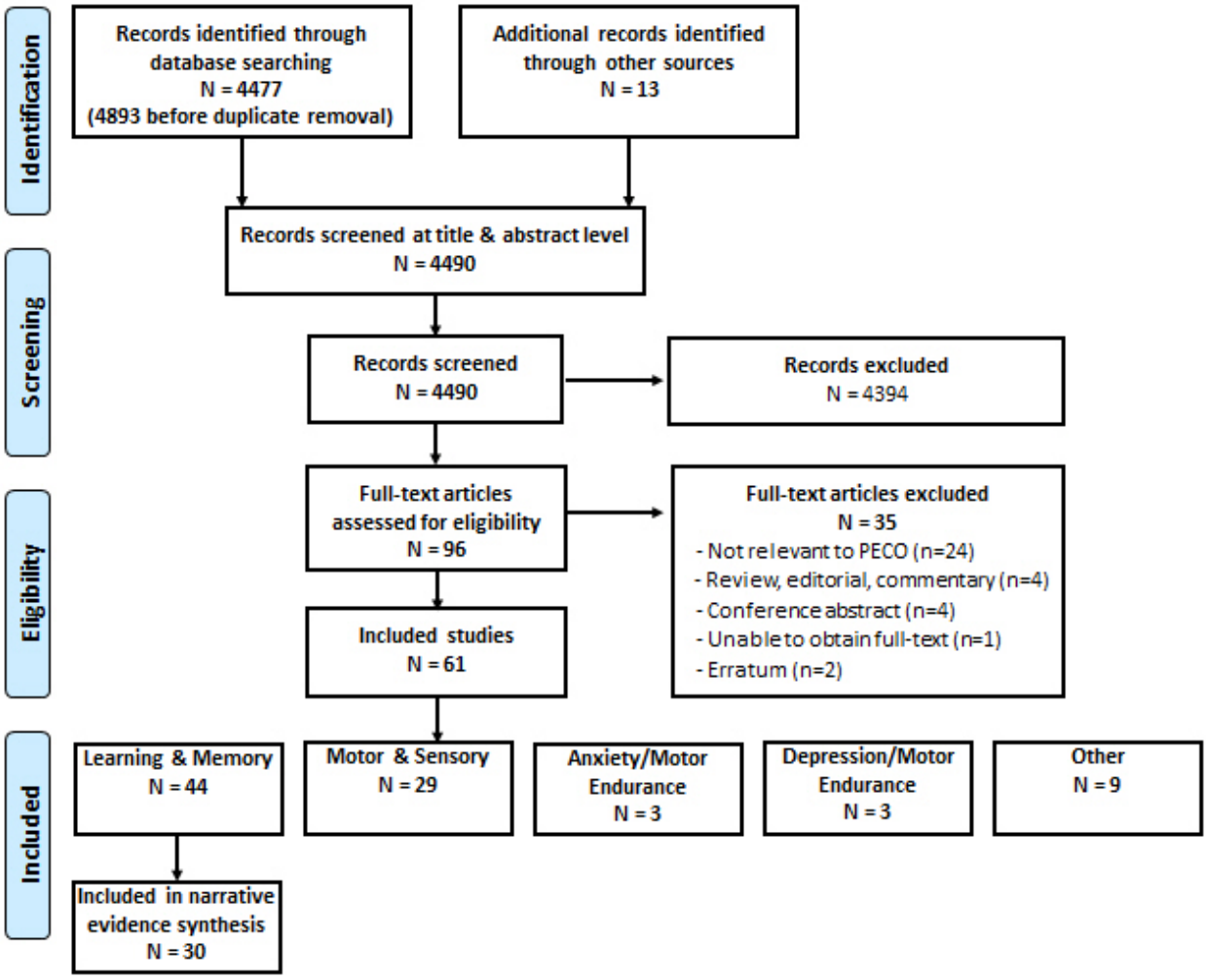


Table S1. Description of relevant studies					
	Learning and Memory	**Motor and Sensory Function**	**Depression/Motor Endurance**	**Anxiety/Motor Activity**	**Other**
Number of studies	44	29	3	3	9
Non-English	15 (34%)	6 (21%)	0 (0%)	0 (0%)	0 (0%)
Species					
rats	35	22	0	1	7
mice	9	7	3	2	2
Life-stage of exposure*					
adult	29	20	3	2	5
developmental	15	10	0	1	5
Doses tested*					
range (ppm, F equivalents)	0.9 - 272	0.9 - 226	0.9 - 90	0.9 - 90	1 - 136
≤10 ppm	17	6	1	1	3
developmental	3	4	--	--	2
11-25 ppm	17	11	0	0	4
developmental	7	6	--	--	4
>25 ppm	29	20	2	2	6
developmental	11	4	--	--	1
Studies with very serious risk of bias					
total	14	9	0	0	3
developmental	9	5	0	0	2
≤10 ppm	3 (100%)	3 (75%)	--	--	2 (100%)
11-25 ppm	4 (57%)	3 (50%)	--	--	2 (50%)
>25 ppm	6 (55%)	2 (50%)	--	--	0
Studies used for evidence synthesis					
total	30				
developmental	6				
≤10 ppm	0				
11-25 ppm	3				
>25 ppm	5				
adult	24				
≤10 ppm	11				
11-25 ppm	9				
>25 ppm	14				

*Numbers may not total because studies often tested multiple dose levels and some studies evaluated effects in multiple lifestages of exposure.