Impact of community water fluoridation on systemic health excluding oral health

An evidence review

November 2022

Kathryn Lambe

Ailish Farragher

Tonya Moloney

Salome Sunday

Jean Long

Disclaimer

Any views expressed in this report are those of the authors and not necessarily those of the Minister for Health, the Department of Health, or the Health Research Board.

Published by:

Health Research Board, Dublin

Copies of this report can be obtained from our website: www.hrb.ie

Cite as: Lambe K, Farragher A, Moloney T, Sunday S, Long J. Impact of community water fluoridation on systemic health excluding oral health: An evidence review. Dublin: Health Research Board; 2022. Available from: https://www.hrb.ie/publications

Health Research Board
Grattan House
67–72 Lower Mount Street
Dublin 2
D02 H638
Ireland

t 353 1 234 5000
f 353 1 661 2335
e hrb@hrb.ie
w www.hrb.ie

Table of Contents

List of Tables

List of Figures

Acknowledgements

We extend our thanks to our colleagues in the Department of Health and expert consultants Professor Jonathan Broadbent and Professor Gary O'Reilly for their valuable contributions to this report. We also thank Joan Quigley for her work on the review protocol. We acknowledge the contributions of our Evidence Centre colleagues Louise Farragher, Caitriona Lee, and Olivia Cagney. Special thanks to our peer reviewers for their feedback, including Dr Richard McNally and Dr Daniel Umpierre.

Abbreviations

Abbreviation	Explanation
ADD	attention deficit disorder
ADHD	attention deficit hyperactivity disorder
ANCOVA	analysis of covariance
ANOVA	analysis of variance
BMD	bone mineral density
BMI	body mass index
CADTH	Canadian Agency for Drugs and Technologies in Health
CENTRAL	Cochrane Central Register of Controlled Trials
CHMS	Canadian Health Measures Survey
CI	confidence interval
CINAHL	Cumulative Index of Nursing and Allied Health Literature
CWF	community water fluoridation
DXA	dual-energy X-ray absorptiometry
EUROCAT	European Registration of Congenital Anomalies and Twins
FSIQ	full scale intelligence quotient
g/cm^2	grams per square centimetre
g/cm^3	grams per cubic centimetre
GEE	generalised estimating equation
GRADE	Grading of Recommendations, Assessment, Development and Evaluations
HOME	Home Observation for Measurement of the Environment
HRB	Health Research Board
ICD-9	International Classification of Diseases, 9th Revision
IQ	intelligence quotient
LILACS	Latin American and Caribbean Health Sciences Literature
MANCOVA	multivariate analysis of covariance
MANOVA	multivariate analysis of variance
MeSH	Medical Subject Headings
mg/L	milligrams per litre
MIREC	Maternal-Infant Research on Environmental Chemicals
mIU/L	milli-international units per litre
NHLBI	National Heart, Lung, and Blood Institute
NHMRC	National Health and Medical Research Council (Australia)
PICO	population, intervention, comparator, and outcome
ppm	parts per million
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
p-value	probability value
SD	standard deviation
SEM	standard error of mean
SIDS	sudden infant death syndrome
TSH	thyroid-stimulating hormone
USA	United States of America
vs.	versus
WHO	World Health Organization
µmol/L	micromole per litre
µmol/mmol	micromole per millimole

Glossary of terms

Term	Explanation
Bias	Bias is a systematic overestimation or underestimation of the association in research. There are many types of bias, such as selection, recall, observer, and interviewer bias. Bias is minimised through good study design and implementation.
Blinding	Blinding is a method used in research to ensure that the people involved in a research study – participants, clinicians, and researchers – do not know which participants are assigned to each study group, or which experienced the exposure or outcome of interest. Blinding is used in order to ensure that knowledge of the type of exposure, treatment, or diagnosis does not affect a participant's response to the treatment, a healthcare provider's behaviour, or an interviewer's approach to data collection.
Case-control study	A case-control study is an analytic observational epidemiological study which examines subjects (cases) with an outcome (disease) back to predetermined exposure (cause), and compares their exposures with self-selected controls that do not have the disease (but are otherwise similar) in order to determine the odds that the exposure may have caused the disease. The odds ratio is the measure of choice in a case-control study. This type of study can be used to identify exposures that cause rare diseases. They contribute low-quality evidence to causality or disease aetiology. The main drawbacks in case-control studies are their potential for recall bias and that they cannot calculate incidence.
Causality	Causality is the relation of cause and effect. The Bradford Hill criteria for causality are: strength of association or effect size; consistency of findings across studies (known as reproducibility); biological credibility (plausibility); specificity (other explanations); a temporal relationship (exposure occurred before the outcome) and biological gradient known as a dose–response relationship; coherence (consistent with other lines of evidence); and analogy (similar agents act similarly).
Chance	Chance is sampling variability which can give rise to a particular result. It is the "luck of the draw". It is an unsystematic over- or underestimation of the cause-and-effect relationship. The *p*-value measures the probability or likelihood that an observed result occurred by chance alone.
Cohort study	A cohort study is a form of longitudinal (analytic observational) epidemiological study in which a group of subjects, called a cohort, is followed over a period of time, and data relating to predetermined exposures and outcomes are collected on two or more occasions over this time period. The incidence (new cases) of the outcome(s) of interest is calculated in the exposed people and compared with the incidence in the non-exposed people. This comparison of incidence is known as relative risk. The data for the cohort can be collected either by following the participants into the future (prospective study) or by asking them about their past (retrospective study). However, retrospective cohort studies are limited by recall bias. One of the indicators of a high-quality cohort study is a loss to follow-up rate of less than 20%. Cohort studies contribute to causality or disease aetiology and provide, at most, moderate-quality evidence.
Community water fluoridation	The practice of artificially fluoridating water with a precise low dose of fluoride as a public health prevention measure to protect teeth from developing caries or cavities. In Ireland, statutory regulations for fluoridation of water supplies stipulate that fluoride may be added to public water supplies, typically in the form of hydrofluorosilicic acid. The 2000 Forum on

Term	Explanation
	Fluoridation recommended that the fluoride level in drinking water in Ireland should be within the range of 0.6 to 0.8 parts per million (ppm), with a target of 0.7 ppm.
Confidence interval	A confidence interval is the range of values (for example, proportions) in which the true value is likely to be found with a degree of certainty (by convention, a 95% degree); that is, the range of values will include the true value 95% of the time.
Confounding	Confounding is when a factor has an association with the exposure and can independently cause the outcome or disease. It can over- or underestimate an effect of interest or association. A confounding variable (also called a confounding factor or confounder) is a variable that has a relationship with both the exposure and the outcome variable. Confounding is controlled for by restricting the study population, matching the study population (for age, sex, geography, and/or socioeconomic factors), randomly selecting the study population, undertaking a stratification in the analysis (for example, by age, sex, geography, and/or socioeconomic factors), and performing regression analysis.
Cross-sectional survey	A cross-sectional survey or prevalence survey is a descriptive epidemiological study in which the presence or absence of both the exposure and outcome is assessed at the same point in time. This study type is vulnerable to the problem of which came first – the exposure or the outcome (likened to the chicken or the egg) – as both exposure and outcome are collected at the same point in time. These types of studies are often used to assess the prevalence of acute or chronic conditions; to inform health planning and evaluation; or to formulate a theory. It can be difficult to control for factors that may be related to the exposure and outcome in cross-sectional studies, so they cannot be used to determine causality. They are sometimes included in the hierarchy of evidence and are considered to provide very low-quality evidence.
Ecological or correlational study	An ecological study is a descriptive epidemiological study carried out using aggregated population-based data to describe a disease (outcome) in relation to a factor of interest (exposure) and is used to formulate a theory (not to prove causality). Both the outcome and exposure are correlated to determine their linear association, which is expressed as a proportion of exposure and outcome that correlate with each other. This study type is vulnerable to 'ecological fallacy', as it is not known whether the individuals who were exposed were the same individuals who experienced the outcome (or disease). These types of studies are not usually included in the hierarchy of evidence and so would only provide very low-quality evidence.
Fluorine	Fluorine is a chemical element with the symbol F and atomic number 9. It is a member of the halogen family. Fluoride is the negative ion of the element fluorine.
Fluorosis	There are two types of fluorosis: Skeletal fluorosis is a bone and joint condition associated with prolonged exposure to high concentrations of fluoride. Skeletal fluorosis is typically seen in regions with high levels of natural fluoride (>3 ppm) in groundwater and may result in severe pain or stiffness in the joints. Dental fluorosis is a tooth enamel defect, which in a mild form is typically observed as mild white lines or opaque white spots on the enamel. Moderate and severe forms of dental fluorosis, which are far less common, cause more extensive enamel changes. More severe forms of dental fluorosis can cause discoloured, pitted, or weakened teeth. As tooth development occurs in the first eight years of life, children are susceptible to fluorosis up to this age. The severe form hardly ever occurs in communities where the level of fluoride in water is less than 2 ppm. Dental fluorosis is caused by children taking in too much fluoride over

Term	Explanation
	a long period when the teeth are forming under the gums. Increases in the occurrence of mostly mild dental fluorosis were recognised as more sources of fluoride became available to prevent tooth decay.
Full Scale IQ (FSIQ)	Full Scale IQ is the most global score of cognitive ability on an IQ test, comprising a combination of scores on a number of subtests.
Generalised estimating equation	The generalised estimating equation is a statistical technique used as a method for modelling longitudinal or clustered data. It is usually used with non-normal data. Unlike mixed-effects or multilevel models, the generalised estimating equation is a marginal model, used to model a population average rather than different parameters for each subject or cluster. It is intended for simple clustering or repeated measures designs.
Guideline value	Guideline values are derived by the World Health Organization for many chemical constituents of drinking water. A guideline value normally represents the concentration of a constituent that does not result in any significant risk to health over a lifetime of consumption. A number of provisional guideline values have been established based on the practical level of treatment performance or analytical achievability. In these cases, the guideline value is higher than the calculated health-based value.
Hierarchy of evidence	The hierarchy of evidence from highest to lowest for primary epidemiological studies is: randomised controlled trials, non-randomised trials, longitudinal cohort studies, case-control studies, and cross-sectional studies. Ecological or correlational studies are not usually on the hierarchy of evidence, as their role is to suggest rather than prove causal relationships.
Incidence	Incidence is a term used to describe the number of new cases of disease or events that develop among a population during a specified time interval.
Index scores	On IQ tests, index scores refer to scores on specific domains of ability, e.g. a test's verbal index (sometimes called verbal IQ) or numeric index (numeric IQ). Index scores are correlated with FSIQ scores but also have some degree of independence from them and from other index scores.
Intelligence Quotient (IQ)	An intelligence quotient (IQ) is a total score derived from a set of standardized tests or subtests designed to assess human intelligence. For modern IQ tests, the raw score is transformed to a normal distribution with mean 100 and standard deviation 15, with reference to a representative national population sample of a relevant age.
Logistic regression	Logistic regression is a statistical technique used in research designs that require the analysis of the relationship of an outcome or dependent variable to one or more predictors or independent variables when the dependent variable is either: (a) dichotomous, having only two categories (for example, whether one uses illicit drugs (no or yes)); (b) unordered polytomous, which is a nominal-scale variable with three or more categories (for example, eye colour (blue, brown, grey, or green)); or (c) ordered polytomous, which is an ordinal-scale variable with three or more categories (for example, the highest level of education completed (none or primary school incomplete, primary school, secondary school, third-level diploma, third-level primary degree, third-level master's, third-level doctorate)).
Mg/L	The unit of measurement for fluoride in water is parts per million (ppm) or milligrams per litre (mg/L). The units are interchangeable; 1 ppm equals 1 mg/L.
Odds ratio	An odds ratio is a statistic that quantifies the strength of the association between two events, A and B. The odds ratio is defined as the ratio of the odds of A in the presence of B and the

Term	Explanation
	odds of A in the absence of B, or equivalently (due to symmetry), the ratio of the odds of B in the presence of A and the odds of B in the absence of A.
Performance IQ	Performance IQ refers to the score on an index/domain on IQ tests measuring a range of non-verbal skills, including fluid reasoning, spatial processing, attentiveness to details, and visual-motor integration.
Ppm	The unit of measurement for fluoride in water is parts per million (ppm) or milligrams per litre (mg/L). The units are interchangeable; 1 ppm equals 1 mg/L.
Prevalence	Prevalence is a term used to describe the proportion of people in a population who have a disease or condition at a specific point in time or during a specific period.
Relative risk or risk ratio	The relative risk or risk ratio is the ratio of the probability of an outcome in an exposed (or intervention) group relative to the probability of an outcome in an unexposed (or control) group and compares the incidence of an outcome in the exposed group with the incidence of the outcome in the unexposed group.
Student's two-tailed unpaired *t*-test	An unpaired *t*-test (also known as an independent *t*-test) is a statistical procedure that compares the means of two independent or unrelated groups in order to determine if there is a significant difference between the means of the two groups.
Verbal IQ	Verbal IQ refers to the score on an index/domain on IQ tests measuring verbal ability.
Z-test	A Z-test is a statistical test used to determine whether two population means are different when the variances are known, and the sample size is large.

Executive summary

Purpose

In 2015, the Health Research Board (HRB) Evidence Centre published an evidence review titled *Health effects of water fluoridation*. The review question in the 2015 publication was "What is the impact, positive and/or negative, on the systemic health of the population (excluding dental health) for those exposed to artificially fluoridated water between 0.4 and 1.5 parts per million (ppm)?".

The 2015 review established that there was no definitive evidence that showed an adverse impact on general health due to water fluoridation at optimal levels (0.4–1.5 ppm). However, it is essential, and required by legislation, to continuously monitor and evaluate the evidence in order to ensure that no new adverse safety issues are present. The 2015 review suggested that the impact of fluoridated water on thyroid function, as well as on bone health (including cancer), required monitoring and reassessment.

Community water fluoridation was introduced in Ireland in 1964 as a public health measure to prevent tooth decay (dental caries); fluoride was added at a level of 1.0 ppm. In 2000, water fluoridation policy in Ireland was the subject of a major review by the Forum on Fluoridation, which recommended that the fluoride level be lowered to a range of 0.6–0.8 ppm, with a target of 0.7 ppm. Ongoing evaluation of the evidence regarding the effects of water fluoridation is extremely important, so that the benefits and potential harms can be accurately weighed by policy-makers.

For these reasons, the 2015 review is now being updated to incorporate the most up-to-date evidence published between 1990 and 2021.

Review question

What is the impact on the systemic health of the human population for those exposed to artificially fluoridated water between 0.4 and 1.5 ppm, compared with non-fluoridated water (less than 0.3 ppm)?

Methods

This review followed the recommended approach to a systematic review, starting with a population, intervention, comparator, and outcome (PICO) question and eligibility criteria. The search concepts were based on water, fluoride, and primary epidemiological studies. The date limitations for the updated systematic review were 1990 to September 2021. The HRB searched five databases and two trial and protocol registers for studies published between 2014 and May 2021. In addition, reference and citation searching and supplemental searches of systematic reviews were completed. The HRB identified primary studies published between 1990 and 2013 from three existing systematic reviews. Only English-language material was included, as due to the time frame of the review, it was not possible to commission translations. In searching the Latin American and Caribbean Health Sciences Literature (LILACS) database, we captured relevant English-language material and ensured a more diverse evidence base. Screening was completed in duplicate for identified titles and abstracts and again for full-text papers. Systematic extraction and quality assessment was completed by one researcher and validated by a second researcher. A feasibility analysis was completed in order to determine if the studies (by study outcome and design) could be combined in meta-analyses, considering population, exposure, comparator, and outcome. Narrative analyses were completed for each outcome. GRADE (Grading of Recommendations, Assessment, Development and Evaluations) of evidence was applied to each outcome, taking account of the risk of bias, inconsistency, indirectness, imprecision, and publication bias.

Findings

A total of 30 studies reported across 37 papers meeting the review eligibility criteria were identified from all stages of the search process. Many studies examined multiple outcomes.

Bone health

Bone characteristics

Eight papers examined the association between fluoridated water and bone characteristics. Four of the eight papers presented data from three cross-sectional surveys, three papers were ecological studies, and one paper was a cross-sectional study that allocated exposure on an ecological or population basis, which should be treated as an ecological study for the purposes of contributing evidence for causality. Overall, the evidence from these studies for the influence of community water fluoridation (CWF) on bone mineral density (BMD) was mixed. Although a number of studies found associations between higher BMD and exposure to water fluoridation in certain skeletal areas (such as the lumbar spine), contradictory findings also exist, and a large number of analyses found no association. Therefore, no theoretical relationship has been firmly established. Additionally, one ecological study found no association between the incidence of osteoporosis and water fluoridation status.

Fractures

Ten papers presented data from 9 studies that examined the association between water fluoridation status and the incidence of a range of fractures, most commonly hip fractures. Seven of the 10 papers were based on ecological studies and 3 papers were based on 2 cross-sectional surveys. Overall, the evidence from these papers for an association between CWF and fracture incidence was mixed, with most analyses pointing to a neutral or, in a few analyses, possible protective effect of fluoridation, although only hip fracture has been extensively studied. It is important to note that none of the included analyses controlled for osteoporosis, which is the leading risk factor for hip fracture.

Neuropsychological outcomes

Seven papers based on four studies examined the association between fluoridated water and neuropsychological outcomes. Four papers based on two studies examined IQ in childhood and adulthood, and one additional paper examined aspects of neuropsychological development in infancy and childhood, which conceptually maps closely to IQ.

The studies investigating the influence of fluoride on IQ and neuropsychological development have mixed findings, variously reporting null, positive, and inverse associations between fluoride exposure and IQ in childhood (however, the small number of positive associations should not be interpreted as evidence for a beneficial effect of fluoride on IQ). Two of the three cohort studies had high loss to follow-up and the one longitudinal cohort study (Maternal-Infant Research on Environmental Chemicals study) has methodological issues that call into question the validity of the findings presented in the papers based upon it. The remaining two papers, based on one cross-sectional study (Canadian Health Measures Survey), present conflicting findings with respect to diagnosis of ADHD, and one analysis demonstrated stronger associations between fluoride exposure and hyperactivity for older youth. A high-quality prospective longitudinal study based on individual-level exposures, taking account of all potential confounding factors, effect modifiers, and cluster design effect, is required in order to strengthen the evidence base on neuropsychological outcomes.

Cancer

Bone cancers

Eleven papers presented data from 10 studies examining the association between bone cancers and fluoridated water status. Eight were ecological papers and three were case-control papers. Osteosarcoma was the cancer examined in 10 papers (9 studies), bone cancers in general were examined by 2 papers, and Ewing sarcoma was examined by 1 paper. The evidence from the ecological studies does not suggest any association between CWF and the incidence of bone cancers. In addition, the relationship between a diagnosis of osteosarcoma and exposure to artificially fluoridated water is unlikely, based on the evidence from case-control studies. Therefore, no relationship can be established.

Other cancers

Two ecological studies examined other cancers. One ecological study examined differences in incidences of any cancer and bladder cancer between fluoridated and non-fluoridated areas, while the other ecological study examined differences in secondary bone cancer incidences according to the extent of fluoridation implementation. The first, an ecological study conducted in England, found a lower incidence of bladder cancer and all cancers in fluoridated areas (compared with non-fluoridated areas) after adjusting for confounders. While these results suggest some possible protective effects of fluoridation against some forms of cancer, that association is based on an ecological study design and cannot be considered causal. In the second ecological study of cancer patients, no relationship was found between the percentage of county-level access to fluoridated water and the prevalence of secondary bone cancer.

Endocrine conditions

Four papers, presenting data from two studies, examined the association between artificially fluoridated water and endocrine conditions. Three papers, based on two studies, examined a range of outcomes related to thyroid functioning, including incidence of diagnoses of thyroid disorders and TSH levels, while one paper examined the incidence of sleep disturbances, which the study authors attribute to the functioning of the pineal gland.

Overall, although the effects of fluoride on thyroid functioning have long been observed, the evidence for an association between CWF specifically and thyroid conditions and outcomes was mixed, and the findings were based on a small number of ecological and cross-sectional studies, not high-quality cohort studies. Therefore, no relationship has been firmly established. A high-quality prospective longitudinal study based on individual-level exposures and taking account of all potential confounding factors and effect modifiers (such as iodine) is required in order to strengthen the evidence base on the relationship between CWF and thyroid or other endocrine conditions. One cross-sectional survey paper generally found no association between fluoride exposure (measured by tap water concentrations and specific gravity-adjusted urinary concentrations) and a range of self-reported sleep outcomes, including sleeping more than the recommended duration, trouble sleeping, and daytime sleepiness, although it found some evidence for a higher risk of sleeping less than the recommended amount with higher fluoride exposure.

Renal conditions

One ecological study monitored the health effects of water fluoridation arrangements in England. The study compared rates of selected non-dental health outcomes (in this case, renal calculi) between areas according to whether the level of fluoride in drinking water was adjusted (fluoridated) or not (non-fluoridated). The study found that the incidence of renal calculi (kidney stones) was lower in fluoridated areas than in non-fluoridated areas, controlling for age, gender, deprivation, and ethnicity. This association is based on an ecological study design and cannot be considered causal.

Birth or birthing abnormalities

Three ecological studies examined birth or birthing abnormalities; the incidence of Down syndrome was the outcome of interest in two of these studies, and the incidences of trisomies, stillbirths, neural tube defects, clefts, and preterm births were each examined by one study. The studies found no association between exposure to fluoridated water and the incidence of Down syndrome, trisomies, neural tube defects, clefts, or stillbirths. One study found that women who received dental cleaning and were exposed to artificially fluoridated water, along with those who received dental cleaning alone, had a significantly lower incidence of preterm births compared with those who had neither, after controlling for confounding variables. This study found that CWF alone had no association with the incidence of preterm births.

Infant abnormalities

One case-control study with ecological assignment of CWF status examined infant abnormalities; in this case, SIDS was the outcome of interest. The study found no association between SIDS and prenatal exposure to fluoridated water. Postnatally, the study also examined the association between SIDS and water fluoridation status and feeding method (breastfeeding compared with formula feeding). No higher risk of SIDS was associated with either breastfeeding or formula feeding in fluoridated areas compared with non-fluoridated areas, nor was there any evidence of an interaction between water fluoridation status and feeding.

All-cause mortality

One ecological study examined all-cause mortality. The study found that the death rate from all recorded causes was lower in fluoridated areas than in non-fluoridated areas, but the effect size was small. This association is based on an ecological study design and cannot be considered causal.

Conclusions

This review, encompassing 30 studies from nine countries, including Ireland, between 1990 and September 2021, indicates that there continues to be no definitive evidence that CWF has negative health effects. We found no conclusive evidence for a link between CWF and most conditions we examined for which research was available, including bone health, cancer, kidney stones, birth and infant abnormalities, and death rates. The evidence is generally of low quality and most of the studies included, due to their designs, cannot provide evidence for any causal relationships.

While bone health and cancer have previously been primary areas of concern for researchers, the findings of this review point to generally mixed or null findings in relation to these outcomes. However, neuropsychological and endocrine outcomes emerged as areas requiring further monitoring. As the existing research in this area is currently limited in scope and interpretation is hampered by methodological problems, further high-quality research is now needed in order to shed light on the impact, if any, of artificial water fluoridation on these aspects of systemic health.

1 Introduction

1.1 Background

Fluorine is a chemical element with the symbol F and atomic number 9. Fluoride is the negative ion of the element fluorine. Any compound, whether it is organic or inorganic, that contains the fluoride ion is also known as a fluoride. Examples include the ionic compounds calcium fluoride (CaF_2) and sodium fluoride (NaF). Ions containing the fluoride ion are similarly called fluorides (e.g. bifluoride (HF_2^-)).

Fluoride is a naturally occurring mineral that is found in varying amounts in surface and groundwater and in some foods. The concentration of fluoride in drinking water is expressed in units of milligrams per litre (mg/L) or parts per million (ppm); these units are equivalent and are used interchangeably by the authors of studies included in this review.

There is considerable variation in the level of naturally occurring fluoride in drinking water around the world, which is largely dependent on geological factors. High levels of naturally occurring fluoride, equal to or exceeding 1.5 ppm, are observed in different parts of the world (see Figure 1); in some areas, well water may contain concentrations of fluoride up to 10 ppm. High levels of naturally occurring fluoride in water occur in approximately 25 countries worldwide. In Asia, high levels of fluoride are observed in parts of India and China. In Latin America, Mexico and Argentina have the highest levels. Parts of North and East Africa also have high levels of fluoride, including parts of Ethiopia, Kenya, Tanzania, Mozambique, Uganda, and Malawi that are transected by the East African Rift Valley.

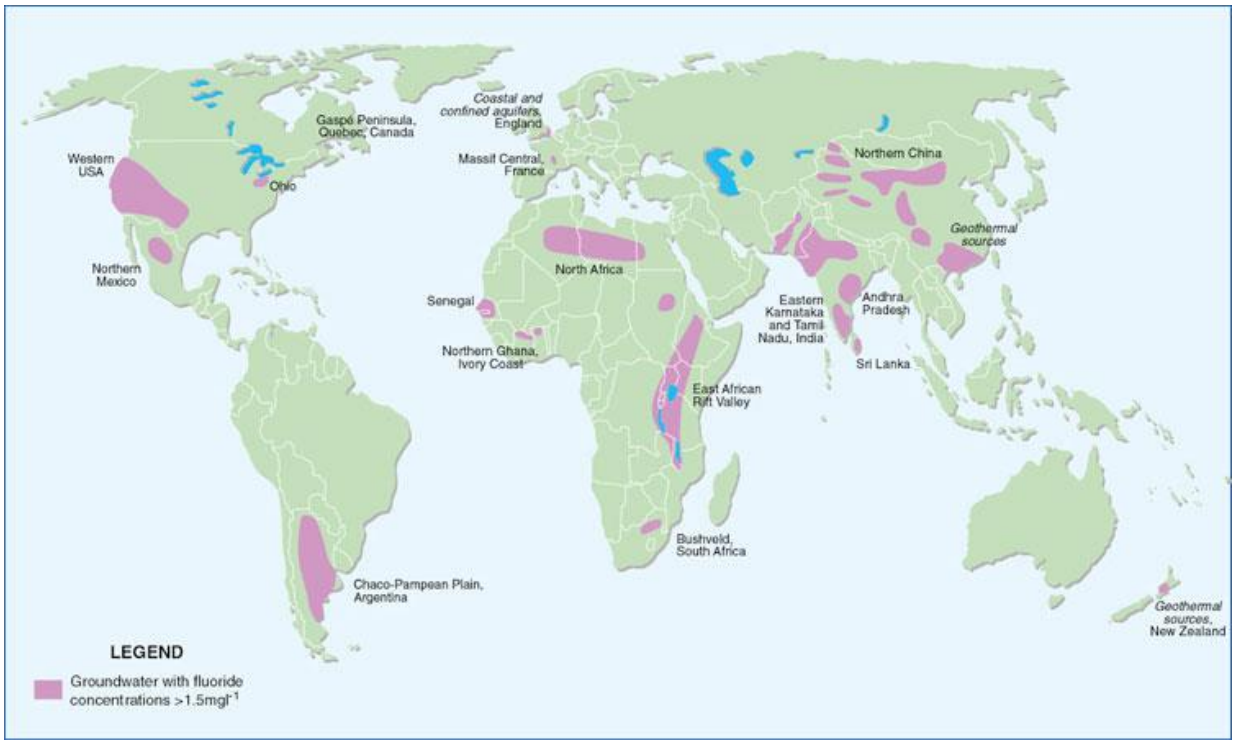

Figure 1 Map of documented occurrences of high fluoride in groundwater (≥1.5 mg/L, equivalent to ≥1.5 ppm)

Source: British Geological Survey, 2021 [1] Reproduced with the permission of the British Geological Survey © UKRI (2021). All Rights Reserved.

1.1.1 Fluoride and health

In humans, fluoride primarily produces effects on skeletal tissues (i.e. bones and teeth). Exposure to fluoride at high concentrations increases the risk of dental fluorosis (pitting or mottling of tooth enamel) and skeletal fluorosis (deposits on bone with adverse changes in bone structure), which may be severe [2]. However, at low concentrations, fluoride in drinking water is also known to have an important protective effect against dental caries, which is a significant public health problem internationally [3–5]. Prevention of caries is an important priority for public health, as it is associated with a reduction in the number of hospital attendances for tooth extractions and anaesthesia, the cost of dental treatment for children, and tooth loss in adulthood [6].

Topical fluoride interventions, including toothpaste and dental products such as varnishes and mouth rinses, also offer considerable benefits for preventing dental caries. These preventative interventions introduce fluoride through direct contact with the exposed surface of the tooth, which increases resistance to decay from bacterial acid attack by inhibiting tooth demineralisation, promoting tooth remineralisation, and inhibiting the activity of bacteria in plaque [6].

However, fluoridated drinking water has the advantage of making fluoride accessible to the entire population of an area, therefore reducing the need for individual compliance and conferring benefits on those who lack access to fluoridated products or treatments or to professional dental care [3,7]. Alternative publicly funded oral health schemes, such as the provision of topical fluoride varnishes through schools, tend to target only high-risk or young populations [6]. Community water fluoridation (CWF) also has particular benefits for reducing caries among children with long-term benefits for oral health. Although fluoride has substantial benefits for caries prevention when delivered topically (e.g. through professional dental treatments) and some of the effect of ingested fluoride may be said to be delivered topically (i.e. through saliva), there is evidence that the preventative effect is maximised by continuous exposure before and after eruption for both adults and children [8]. Other methods to deliver systemic fluoride include milk, salt, and supplements, but these are not of interest to the Department of Health in Ireland, as it has chosen to deliver systemic fluoride through CWF.

Some countries therefore control fluoride levels in the public water supply by artificially supplementing or removing fluoride in order to reach an optimal level or range, while remaining within the World Health Organization (WHO) guideline value of 1.5 ppm, which was set in 1984. The WHO states that a guideline value "represents the concentration of a chemical constituent that does not result in any significant risk to the health of the consumer over a lifetime of consumption" [9] p6. However, it also states that "guideline values are not formal standards or regulatory limits and are not to be taken as strict limits such as 'maximum permissible concentrations'" [9] p5. Rather, they are intended to provide quantitative risk assessment information for regulatory authorities to make decisions in prescribing limits and standards in relation to environmental exposures (e.g. water and air quality) for a specific population. As such, the exact meaning of the guideline value for fluoride has been framed in a number of ways by various publications, stating that maximum concentrations should fall within this level or that levels above 1.5 ppm carry increased health risks [2,10].

Recommended levels for artificially fluoridated water are usually between 0.5 and 1.0 ppm [2]. The amount of fluoride in drinking water considered to be optimal varies regionally; recommendations must take into account factors such as average daily water consumption (which may be higher in hotter climates) and availability of fluoride from other sources, such as air, food, tea, and dental products [10]. Table 1 presents an overview of the effects of fluoridated water at various concentrations on skeletal tissues.

Table 1 Effects of fluoridated water at various concentrations on skeletal tissues

Fluoride levels in water	Effects on skeletal tissues
0.0–0.3 ppm	Unlikely to confer benefits to dental health; increased risk of caries [11,12]
0.5–1.0 ppm	Recommended level for artificially fluoridated water supplies (varies according to local environmental factors, including climate), providing protection against dental caries, tooth decay, and tooth loss for children and adults; increased risk of mild dental fluorosis [2]
≥1.5 ppm	Increased risk of moderate or severe dental fluorosis [2]
3.0–6.0 ppm	Increased risk of skeletal fluorosis [2]
>10.0 ppm	Increased risk of crippling skeletal fluorosis [2]

1.1.2 Community water fluoridation

Water fluoridation is usually accomplished by adding NaF, fluorosilicic acid (H_2SiF_6), or sodium fluorosilicate (Na_2SiF_6) to drinking water in which the natural fluoride concentration is sub-optimal. The practice began in 1945, when Grand Rapids, Michigan, in the United States of America (USA), became the first city in the world to artificially fluoridate its drinking water, following results of epidemiological studies showing a link between raised levels of fluoride in drinking water and reduced prevalence and severity of tooth decay in local populations [13]. Fluoride may also be removed from water with too high a concentration to adjust to optimal levels. Fluoridation is rarely performed in isolation and usually carried out as part of a water treatment process, including coagulation, flocculation, filtration, chlorination, fluoridation, and pH correction.

The estimated number of people with access to artificially fluoridated water worldwide as of November 2012 was 377,655,000 in 25 countries, including Argentina, Australia, Brazil, Brunei, Canada, Chile, Fiji, Guatemala, Guyana, Ireland, Israel (ceased in 2014), Libya, Malaysia, New Zealand, Panama, Papua New Guinea, Peru, Serbia, Singapore, South Korea, Spain, the United Kingdom, the USA, and Vietnam. In 2012, these countries also had an estimated 17,910,000 people with access to naturally fluoridated water at or around the optimal level (i.e. 0.5–1.0 ppm), bringing the total number of people with access to optimally fluoridated water in those countries to 395,565,000 [14]. Estimates from 2020 of the proportion of populations in countries worldwide receiving government-regulated fluoridated water are shown in Figure 2.

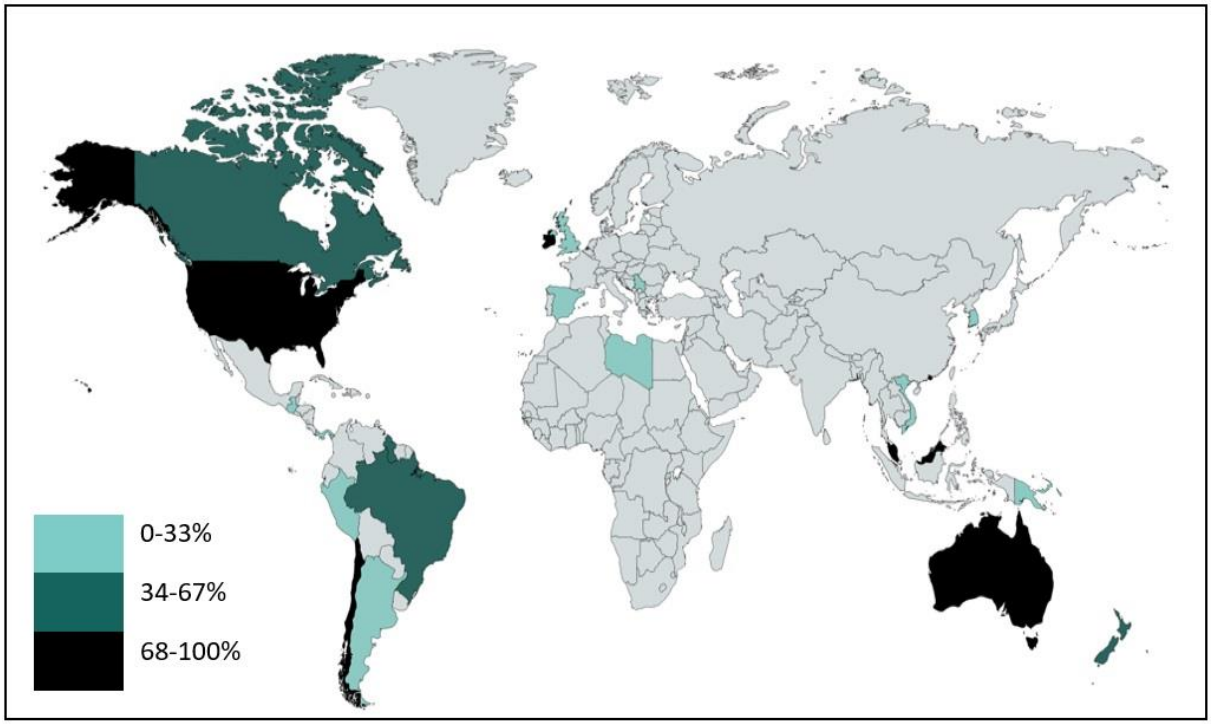

Figure 2 Proportion of the population receiving government-regulated fluoridated water

Source: Johnston and Strobel, 2020 [15]

1.1.3 Community water fluoridation in Ireland

CWF was introduced in Ireland in 1964 following the Health (Fluoridation of Water Supplies) Act, 1960 [16]; fluoride was added at a level of 1.0 ppm. In 2000, water fluoridation policy in Ireland was the subject of a major review by the Forum on Fluoridation, which was established by the then Minister for Health and Children. In light of both international and Irish research showing an increasing occurrence of dental fluorosis [17], the Forum on Fluoridation recommended that the fluoride level in drinking water be lowered from 1.0 ppm to a range of 0.6–0.8 ppm, with a target of 0.7 ppm [18]. This policy was implemented in 2007 [19].

The Fluoridation of Water Supplies Regulations 2007 [19] stipulate that fluoride may be added to public water supplies either in the form of hydrofluorosilicic acid, or in such other form as may be approved by the Minister for Health and Children. It is further stipulated that the fluoride content of public water supplies to which fluoride has been added shall be determined daily at the water treatment plant. Water supplied by local government (which services all urban areas) is required to be fluoridated; however, private water supplies from wells or local community "group schemes" are not required to be fluoridated [20]. In 2017, just over 71% of people living in Ireland had access to publicly provided CWF at an average annual cost to the State of €2.15 per capita of population receiving fluoridated water [21].

1.1.4 Policy considerations

Public health policies should be based on sound scientific evidence about risks and benefits, and on an economic evaluation of interventions to address a specific issue in a population. Decision-makers should also be cognisant of the impact of not employing a proven intervention.

CWF is a cost-efficient intervention that can reach large populations without necessitating the active participation of individuals, and it can deliver oral health benefits to a broad spectrum of people, reducing

disparities in oral health [5,22]. In Ireland, despite current access to numerous fluoride sources and a reported increase in the prevalence of enamel fluorosis, CWF remains a cost-effective public health intervention for Irish schoolchildren [21]. However, there is opposition to, and scepticism regarding, the practice of artificially fluoridating water supplies, both in Ireland and internationally.

Arguments against CWF include concerns about negative environmental impacts [23] and the ethics of the practice. While CWF is implemented with the goal of reducing inequalities in dental health by providing benefits to all, regardless of age, socioeconomic status, or access to dental care, the fact that it is a mass intervention removes individual choice and raises difficult questions about the right to refuse health interventions [13]. Disagreement about the quality of the evidence base regarding benefits and harms [13], and about the accuracy with which this evidence is represented on both sides of the debate [24], has created a tense discourse around CWF in the public sphere.

Concerns have been raised about potential harmful side-effects of fluoridation (e.g. dental and skeletal fluorosis, bone cancer, disruption to thyroid function, and neuropsychological and neurodevelopmental effects). The scientific evidence to date indicates that negative health effects are improbable but cannot be ruled out completely (see Appendix A for an overview of the theoretical basis for links between fluoride exposure and a range of health outcomes). It must be noted that many of the concerns about the adverse health effects of fluoride result from findings in regions with very high fluoride levels in naturally fluoridated water (1.5–10.0 ppm), which are 2–12 times higher than the levels of fluoride in the water in Ireland (0.6–0.8 ppm). Therefore, results from these studies cannot be equated with the situation in Ireland. In fact, studies in regions with high levels of naturally occurring fluoride frequently compare the health of people with very high levels of fluoride in their drinking water (>1.5 ppm) with that of a comparison group living in a nearby area with naturally occurring fluoride within WHO-permissible limits (≤1.5 ppm). The comparison group's exposure to fluoride in areas with naturally fluoridated water, although sometimes higher than the exposure for people who drink artificially fluoridated water, is classified by researchers as normal exposure or as low risk for fluoride-related health effects [25]. The potential effects of artificially fluoridated water on systemic health (excluding oral health) are the primary focus of this report, which will examine the evidence base underlying these concerns. Oral health outcomes relating to CWF will be addressed in a sister Health Research Board (HRB) publication for publication in 2023.

1.2 Review objectives

In 2015, the HRB Evidence Centre published an evidence review titled *Health effects of water fluoridation* [25]. The review question in the 2015 publication was "What is the impact, positive and/or negative, on the systemic health of the population (excluding dental health) for those exposed to artificially fluoridated water between 0.4 and 1.5 ppm?"[25] p19.

The 2015 review established that there was no definitive evidence that showed an adverse impact on general health due to water fluoridation at optimal levels (0.4–1.5 ppm). However, it is essential, and required by legislation, to continuously monitor and evaluate the evidence in order to ensure that no new adverse safety issues are present. The 2015 review suggested that the impact of fluoridated water on thyroid function, as well as on bone health (including cancer), required monitoring and reassessment.

As mentioned in Section 1.1.3, following the publication of the findings of the Forum on Fluoridation in 2002, the fluoride concentration in Irish water supplies was reduced to 0.7 ppm. This was done in order to combat the increased prevalence of dental fluorosis among children and adolescents [17], while maintaining the advantages of the protective effects of fluoride against caries. However, the reduction of fluoride levels in drinking water has not reduced the levels of fluorosis, and the difference in dental decay

prevalence among 12-year-olds between fluoridated and non-fluoridated communities is less substantial than in previous years [26]. While evidence from other countries identifies the inadvertent ingestion of fluoride toothpaste as the more likely cause of fluorosis among children and adolescents, this nonetheless underscores the importance of ongoing evaluation of the evidence regarding the effects of water fluoridation so that the benefits and potential harms can be accurately weighed by policy-makers.

For these reasons, the 2015 review is now being updated to incorporate the most up-to-date evidence. The review takes the form of a traditional systematic review of primary quantitative epidemiological studies, with results synthesised by outcomes and presented in the context of existing knowledge about potential mechanisms of action for fluoride.

1.3 Review question

The review question has been modified to include artificially fluoridated water only and exclude optimally and excessive naturally occurring water fluoridation. The question is:

What is the impact on the systemic health of the human population, excluding oral health, for those exposed to artificially fluoridated water between 0.4 and 1.5 ppm compared with non-fluoridated water (less than 0.3 ppm)?

2 Methods

2.1 Review design

A systematic review was the preferred type of evidence synthesis for this research question, as the intention was to systematically gather and synthesise existing evidence in order to provide an up-to-date summary of the state of research knowledge on the intervention of interest (i.e. CWF), with a view to carrying out a quantitative synthesis through meta-analysis if feasible.

A number of other types of evidence synthesis were considered and ruled out, for the following reasons [27]:

- **Individual patient data meta-analysis** was not considered feasible, given the preponderance of ecological studies in this field.
- As one of the goals of the project was to carry out meta-analyses of individual studies, and the review question was relatively specific and limited only to studies of areas with artificially fluoridated water or non-fluoridated water (whereas many reviews include areas with naturally fluoridated water within acceptable levels), an **umbrella review** was not considered appropriate.
- As the goals of a **scoping review** are to determine the size and nature of the evidence base for a given topic area and identify gaps in the research in service of future primary research or a full systematic review, this type of review would not have served the purposes of the current project and so was not considered appropriate.
- Similarly, a **rapid review** was deemed inappropriate, as there was no requirement to complete the project within a particularly tight time frame, which would have necessitated compromises in the comprehensiveness of the search and synthesis.

For these reasons, a standard systematic review design was used to answer the research question [27]. Published studies and other materials were sourced via systematic database searches and supplemental searches. Only primary research studies have been included.

The review is presented here in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) checklist [28] (see Appendix B for checklist). The study protocol was registered and is available to view on PROSPERO (ID CRD42021269654).

2.2 Comparison of approaches: 2015 and 2022 reviews

This review is an update of the HRB's 2015 review of a similar research question. There are some methodological differences between the 2015 review and this 2022 update. The rationale for these changes is explained below.

Focus on primary studies: The 2015 review included the results of systematic reviews alongside results from primary studies. Only primary studies have been included in this 2022 update. This was to allow for the extraction of data from individual primary studies, which could then be synthesised in a meta-analysis where appropriate. The reference lists of the 2015 review and the systematic reviews [5,22] that were included in it were screened, and any primary studies meeting the inclusion criteria for this update review were selected for inclusion and extracted.

Date limit: The inclusion criteria of the 2015 review imposed no date limit. In this update review, only studies published from 1990 to May 2021 were considered for inclusion in order to capture research that

has been carried out using modern, up-to-date methods for sampling, measurement of exposure and outcomes, and diagnosis.

Database selection: The 2015 review search strategy included database searches of MEDLINE, Embase, the Cochrane Central Register of Controlled Trials (CENTRAL), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and PsycINFO. In this update review, we did not search CINAHL or PsycINFO and instead included the Latin American and Caribbean Health Sciences Literature (LILACS) database. This was done on the recommendation of the information specialist, given the strict time constraints for the review and the substantial overlap of records that CINAHL has with PsycINFO and MEDLINE. This follows the most favourable database combinations for literature searches to be included in systematic reviews [29]. We are confident that the combination of databases and supplemental search strategies (outlined in Section 2.4) have captured all the relevant literature.

Exclusion of areas with naturally occurring fluoride: While the 2015 review included studies of areas with natural fluoridation within and above recommended levels, these studies are excluded from the update. The reasons for this are threefold:

1. Many areas with naturally occurring fluoride in drinking water report fluoride concentration levels far exceeding the 1.5 ppm WHO guideline level. Exposure to fluoride at this level is not a useful reference point for policy decisions being taken in Ireland.
2. These areas may also have other toxic materials (e.g. heavy metals) in the water, and so the effect of naturally occurring fluoride cannot be assessed in isolation.
3. In other areas, naturally occurring fluoride in drinking water falls within the 1.5 ppm WHO guideline level (i.e. at similar levels to those provided under CWF), including parts of the USA. However, in the Irish context, only artificially fluoridated water is of interest, and so areas with artificial fluoridation are the most useful reference point for decisions being taken in Ireland.
4. Monitoring of fluoride levels may not be seen by public health officials as a necessary use of scarce resources in areas with naturally occurring fluoride within the WHO guideline levels.

For these reasons, the exposure of interest was limited only to artificially fluoridated water, excluding areas with naturally occurring fluoride in drinking water within or above recommended levels. The comparator was limited only to non-fluoridated water (generally considered to have a fluoride concentration of <0.3 ppm); studies that compared artificial fluoridation with naturally occurring fluoride levels above 0.3 ppm were excluded. We believe that this restricted definition of the exposure of interest allows for a much more specific and more appropriate analysis to inform Irish policy decisions.

The refined eligibility criteria and search strategy for this update are described in full detail in Sections 2.3 and 2.4, respectively.

2.3 Eligibility criteria

The search strategy was prepared and studies were screened for inclusion on consideration of the eligibility criteria indicated in Table 2.

Regarding the comparator, no strict criteria or cut-off exist to define non-fluoridated water. However, in empirical studies of the effects of fluoridation, 0.3 ppm or below is commonly used as a cut-off to designate low/negligible concentrations or non-fluoridated water [11,30–32]. It is considered unlikely that fluoride levels below 0.3 ppm confer benefits to dental health [11,12]. In order to be eligible for inclusion, studies must include areas with non-fluoridated water as a comparator. However, it was not necessary for the analysis to compare non-fluoridated and fluoridated areas dichotomously; studies using continuous

measures of fluoride exposure (including fluoride concentrations in tap water and urinary fluoride) that include both areas with and without CWF were also eligible for inclusion. Studies measuring urinary fluoride as a metric of exposure were also included. Urinary fluoride is a measure of total fluoride intake, not only fluoride intake from tap water. However, other sources of fluoride are not likely to differ across CWF and non-CWF areas.

The outcomes listed were identified in the 2015 evidence review by the HRB [25] and in a 2019 review by the Canadian Agency for Drugs and Technologies in Health [6], with the exception of sleep disorders; this outcome was added later as it emerged during scoping.

Table 2 Eligibility criteria

Criterion	Inclusion	Exclusion
Population	Human populations of any age	Animals
Intervention or exposure	Artificially fluoridated water (fluoride level 0.4–1.5 ppm)	Intervention areas with naturally occurring fluoride >0.3ppm Mixed artificially fluoridated water and naturally fluoridated water where data cannot be separated
Comparators	Non-fluoridated water (natural level of fluoride level <0.3 ppm)	No comparator
Outcomes	• Bone health: bone mass or mineral density, skeletal fluorosis, fractures, bone cancer, osteosarcoma, osteoporosis • Neuropsychological: intelligence quotient (IQ) and cognitive function, neurotoxicity, Alzheimer's disease, autism spectrum disorder • Cancer: all cancers, bone cancer, bladder cancer • Cardiovascular disease: hypertension, atherosclerosis, myocardial infarction • Kidney disease • Thyroid disorders: goitre, hypothyroidism, hyperthyroidism • Diabetes • Immune system disorders • Sleep disorders • Congenital abnormalities • Fertility, miscarriage, stillbirth, and preterm or premature births • All-cause mortality	Oral health (without presence of any eligible outcomes, e.g. cancer)
Study design	Primary quantitative or epidemiological study designs: • Randomised controlled trials • Controlled clinical trials • Retrospective/prospective cohort studies • Case-control studies • Cross-sectional surveys • Ecological/correlational studies	Case studies Opinion pieces/editorials Qualitative studies Reviews Conference abstracts
Date range	1990–May 2021	Pre-1990

2.4 Identifying research evidence

2.4.1 Evidence from 1990 to 2013

Studies published between 1990 and 2013 were sourced from three systematic reviews: the original 2015 HRB review; a 2000 review by McDonagh *et al*., commonly known as the York review [5]; and a 2007 Australian National Health and Medical Research Council (NHMRC) review [22]. The titles and abstracts of the studies included in each of these reviews (n=389) were screened against the eligibility criteria by one reviewer (either KL or TM). Full-text papers from the screening on title and abstract phase were then sourced (n=26) and read closely, using the same inclusion/exclusion criteria, by two reviewers (either TM and JL or KL and JL). Disagreements were resolved by further review until consensus was reached for each item. Reasons for exclusion were recorded for any excluded papers (see Appendix C).

2.4.2 Evidence from 2014 to 2021

After discussion with the review team, the approach to identifying research evidence was undertaken with reference to the inclusion criteria outlined in Table 2. Systematic reviews were identified during the scoping phase, but only primary quantitative studies were considered for review by the research team and used in the final analysis.

A systematic and comprehensive search of appropriate databases was carried out and was supplemented by a search of registers, grey literature, and repositories of clinical trials and systematic reviews. At the end of the extraction process, a brief date-specific search of the databases was undertaken in order to capture any relevant material which may have been newly published since the initial database searches in May 2021.

Materials retrieved from the databases and other searches were deduplicated and screened, and reference/citation chasing was carried out on included articles. Four screeners (KL, AF, TM, JL) were involved in the two-stage process of screening on title/abstract and screening on full text.

2.4.2.1 Search concepts

The search strategy emerged from a population, intervention, comparator, and outcome (PICO) framing of the research question and was based around the concepts of artificially fluoridated water (intervention) within the human population (population), with non-fluoridated water as the comparator. The systemic health outcomes were not included in the search strategy in order to allow for the inclusion of new outcomes. The search was therefore based around three concepts: water, fluoride, and primary quantitative studies (see Figure 3).

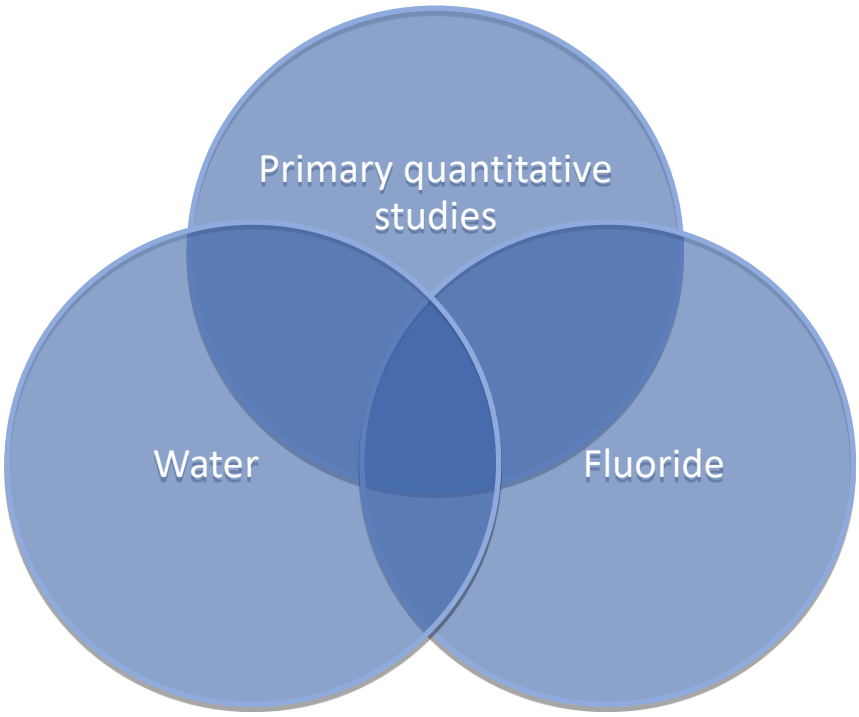

Figure 3 Graphical representation of search concepts

2.4.2.2 Scoping

The initial scoping search was carried out in Ovid MEDLINE, the Cochrane Central Register of Controlled Trials (CENTRAL), Epistemonikos, and Google, with terminology based on the PICO. The 2015 HRB review [25], the 2000 York review [5], and the 2007 NHMRC review [22] were retrieved at this time and were treated as core reviews. The references of those reviews were also included as scoping material. The initial scoping search informed the language of the comprehensive systematic search of selected databases and other resources.

2.4.2.3 Search resources

The selection criteria for reputable sources of evidence were carefully considered in order to ensure the retrieval of a broad range of published clinical and pharmacochemical materials across a wide range of relevant journals and geographic areas. Ovid MEDLINE and Ovid Embase were selected on this basis. The LILACS repository was also selected to address any unintended bias towards European and North American research, as well as to reflect the availability of published evidence on fluoridated water in South America. The Cochrane Database of Systematic Reviews and the CENTRAL database were also searched. The choice of databases was informed by the recommendations in Chapter 4 of the *Cochrane Handbook for Systematic Reviews of Interventions* [33]:

> *The Cochrane Central Register of Controlled Trials (CENTRAL) and MEDLINE, together with Embase (if access to Embase is available to the review team) should be searched for all Cochrane Reviews. Additionally, for all Cochrane Reviews, the Specialized Register of the relevant Cochrane Review Groups should be searched, either internally within the Review Group or via CENTRAL.* [33] p67

The initial scoping search was carried out in Ovid MEDLINE, and was then translated for use in Google, CENTRAL, and Epistemonikos (a database of systematic reviews and primary research that sources material from 26 other databases) with terminology based on the PICO [28].

2.4.2.4 Search strategy

The search strategy emerged from a PICO framing of the research question and was based around the concepts of artificially fluoridated water (intervention) and its impact on systemic health – excluding oral health – (outcome) within the human population (population), compared with non-fluoridated water as the comparator.

Using terms that would capture evidence while excluding non-relevant material proved challenging. Search strategies of systematic reviews on similar themes were reviewed and the validity of several trial searches was assessed to inform the approach. The following approach was taken: the search strategy used detailed search terms around the 'fluoride' concept (such as multiple spellings of 'fluoride' and derived concepts) as well as related Medical Subject Headings (MeSH) terms (e.g. exp *Fluoride/or exp Fluorides/or exp Fluoridation/or exp Fluorine/*). In addition, the search utilised broad language around the concept of water (e.g. exp *Water/or exp Water Supply/*and *water.mp*). This captured the widest range of material relating to concepts such as community water, water treatment, etc. The systemic health (excluding oral health) concept was addressed in the block of MeSH and free text terms used to identify quantitative research [34]. Again, this broad treatment ensured that we did not exclude any relevant systemic health evidence from our search results. Terms limiting the search by date (2014 to May 2021) and population (human) and excluding letters, editorials, and newspaper articles were also included.

The strategy was designed by two information specialists on the review team (Ailish Farragher (AF) and Louise Farragher) and informally peer reviewed by a third information specialist (Caitriona Lee) using the headings of the Peer Review of Electronic Search Strategies checklist (outlined in the *PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration* document) [35].

The complete search strategies are presented in Appendix D. Searches of all databases were carried out on 19 May 2021, with the exception of LILACS; this search was carried out on 24 May 2021.

Details of the supplementary grey literature search are presented in Appendix E.

2.4.2.5 Screening

As the number of results retrieved for screening was substantial (n=3,259), the screening was divided into two teams of two screeners (KL and JL; TM and AF) using EPPI-Reviewer [36]. Double screening was done after the results were divided into two batches, and each item was screened by two screeners on title and abstract. The initial screening criteria (e.g. exclude on duplicate, exclude on population, exclude on intervention, etc.) were based on the agreed review PICO. At the end of the double-screening process, any disagreements were resolved by both screeners further reviewing the articles in question against the inclusion/exclusion criteria until a consensus was reached. The rate of agreement was 96.8%, pooled for the two teams of screeners. Any study without an abstract was sourced at this stage and a decision was made regarding whether to include or exclude it for full-text screening. Where duplicate items were identified, one record was marked in the EPPI-Reviewer database as a duplicate and was excluded.

Full-text papers from the screening on title and abstract phase were then sourced (n=191) and read closely, using the same inclusion/exclusion criteria, in EPPI-Reviewer by each team. Again, double screening was done after the papers were divided into two batches, and each item was screened by two screeners. No disagreements on inclusion/exclusion arose. Reasons for exclusion were recorded for any excluded papers (see Appendix C).

2.4.2.6 Reference and citation searching

The process of citation/reference chasing was carried out on the papers selected for inclusion (n=35) using the complete reference lists published in each paper. Citations and references were sourced using

the online web tool citationchaser [37], as well as the online resources PubMed, ResearchGate, and various publisher websites. All references and forward citations were entered into EndNote library and deduplicated; 1,610 results remained after deduplication. Previously screened or previously included papers were removed, as were papers that were completely out of the scope of the review. This was done in EPPI-Reviewer by the information specialist (AF) using the eligibility criteria. The final results (n=14) underwent full-text screening in EPPI-Reviewer by two members of the review team (KL, JL) for inclusion and one paper was selected to be included in the final analysis (n=36).

2.4.2.7 Final searches

A brief search, revisiting the Ovid Embase and Ovid MEDLINE searches with a date limit of 2021 to 28 September 2021, was carried out on 28 September 2021. The 535 results were screened on title and abstract by one team member (AF). Of the 28 results identified for full-text screening, which were then double-screened by two members of the review team (KL, JL), no further relevant results were identified for inclusion.

One paper [38] was brought to the attention of the research team by a consultant on the project. The paper was accepted for publication after the final searches were complete; however, the study met the inclusion criteria and presented data on neuropsychological outcomes, an area in which high-quality evidence is lacking. For these reasons, the paper was included in the review, bringing the total number of papers included in the final analysis to 37.

2.5 Data extraction

Data were extracted by a single reviewer (KL) into a bespoke extraction sheet in Microsoft Excel. Data on the following parameters were extracted: study author(s), year of publication, research question, primary study design, study country, length of study period, study exposure(s) or cases for case-control study, length of exposure, study comparator(s), study outcome(s), sample size recruited, sample size for analysis, mean age in years, gender (specifically, the proportion of females in the sample), and detailed results. Journal websites for the included articles were checked for supplementary data and errata. Verbatim extraction was completed where feasible and care was taken when extracting numeric results. Where multiple time points, measures, or analyses were presented, all results that were compatible with each outcome domain in each study were extracted. Where information was missing, unclear, or conflicting, this was noted and a conservative approach was taken to any interpretations of conflicting information. Where possible, values were computed based on presented data in order to allow for easier comparison across studies (e.g. conversion of incidence per 100,000 to incidence per 1,000,000). In relation to the interchangeable units ppm and mg/L, we have used the units used by the original study authors in each case.

Extracted data were verified independently by a second reviewer (JL) against a clean copy of the publication. There were no text errors and few numeric errors. Further details of the extraction form are available in Appendix F.

Once extracted, the papers were organised by outcome and then by study design, taking account of the hierarchy of evidence. Our approach is informed by the taxonomy of study designs defined by Hennekens and Buring (1987) [39] and by the NHMRC's [22] hierarchy of evidence, which provides a ranking of study designs considered most relevant for providing information about aetiology and harms. The hierarchy (from highest to lowest) for primary aetiology studies is: longitudinal cohort studies, case-control studies, and cross-sectional studies (see Table 3).

Table 3 Hierarchy of evidence for aetiology/harms

Level	Study design	Notes
I	A systematic review of level II studies	A systematic review will only be assigned a level of evidence as high as the studies it contains, excepting where those studies are of level II evidence
II	A prospective cohort study	
III-1	All or none	All or none of the people with the risk factor(s) experience the outcome. For example, no smallpox develops in the absence of a specific virus, and clear proof of the causal link has come from the disappearance of smallpox after large-scale vaccination.
III-2	A retrospective cohort study	
III-3	A case-control study	
IV	A cross-sectional study	

Source: National Health and Medical Research Council, 2007 [22]

Ecological or correlational studies are not usually on the hierarchy of evidence, as their role is to suggest rather than prove causal relationships, and features of their design usually preclude the ability to test epidemiological hypotheses [39]. As data for ecological studies is collected at population level, it is not possible to link exposure with outcomes for specific individuals, nor is it possible to control for confounding variables. A description of each of the study types (ecological or correlational study, cross-sectional survey, cohort study, and case-control study) may be found in the Glossary of terms.

Where exposure or non-exposure to artificially fluoridated drinking water was assigned to participants based on the prevalence of CWF in the region in which data were collected, rather than the CWF status of their specific residential address or history of residential addresses, these studies were described as having allocated exposure on an ecological/population basis. Therefore, these study designs were reduced to the level of ecological designs for the purposes of contributing evidence for causality.

2.6 Quality assessment

2.6.1 Ecological, cross-sectional, and cohort studies

Ecological, cross-sectional, and cohort studies were assessed for methodological quality by a single reviewer (JL) using the National Heart, Lung, and Blood Institute's (NHLBI's) quality assessment tool for observational cohort and cross-sectional studies [40]. This 14-item tool is designed to include both cohort and cross-sectional studies, which were the most frequently cited study designs in the 2015 HRB review. While ecological studies were also frequently cited in the 2015 review, we were unaware of any tool that would specifically address this type of study design, so they were assessed using the NHLBI's cohort and cross-sectional study tool. The reviewer extracted verbatim text from each paper in order to support the reason for each score.

The tool is presented in full in Appendix G. For each study, an overall quality rating was calculated using a bespoke system, based on essential criteria for high-quality cohort and cross-sectional studies [39]. Five items from the tool were selected and scored as outlined in Table 4. The items chosen identified aspects of studies that were most likely to introduce bias to the results through unrepresentative sampling, loss to follow-up, and confounding.

Table 4 Overall quality rating calculation for ecological, cross-sectional, and cohort studies

Item	Scoring
3. Was the participation rate of eligible persons at least 50%?	Yes: 1.0 No: 0.0
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes: 1.0 No: 0.0
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes: 1.0 Partly: 0.5 No: 0.0
13. Was loss to follow-up after baseline 20% or less?	Yes: 1.0 No: 0.0
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Almost all: 1.0 Partial: 0.5 Limited: 0.0 None: 0.0

Note: Responses of "Not reported" and "Not applicable" were scored 0.0 for each item. For item 14, key potential confounding variables were identified based on established risk factors for the condition under consideration (see Appendix A and Table 52); while some studies controlled for a large number of variables in their models, only these key confounding variables were considered for item 14.

For each study, the scores were summed (for a total score ranging from 0.0 to 5.0). Studies scoring less than 3.0 were rated 'low quality', studies scoring 3.0 were rated 'moderate quality', and studies scoring 3.5 or more were rated 'high quality'. As many studies were cross-sectional in nature (point-in-time surveys) and scored 0.0 on item 13 (loss to follow-up not applicable), the maximum possible score for these studies was effectively capped at 4.0; for this reason, the threshold for 'high quality' was set at 3.5, rather than 4.0, in order to allow more effective differentiation of studies at the upper end of the range of scores.

Some studies examined multiple outcomes; for example, the ecological study by Young *et al.* [41] examined a large number of health outcomes across the areas of bone health, cancer, renal conditions, birth or birthing abnormalities, and all-cause mortality. As the quality of factors such as control for confounding, definition of exposure and outcome measures, and sample size can vary across outcomes even within the same study, we therefore chose to carry out the quality assessment independently for each outcome for a number of studies. For example, for studies examining the impact of water fluoridation on cancer (see Section 3.5), three separate quality assessments were carried out for the Young *et al.* (2015) paper in order to describe the quality of evidence related to impacts on osteosarcoma, bladder cancer, and all cancers.

Quality assessment ratings were verified independently by a second reviewer (KL) against a clean copy of each publication. The main area of difference between the reviewers was with respect to the scoring of the quality of control for confounding. It was therefore decided that the review team would identify confounders in the medical epidemiological literature for each outcome of interest and compare these with the confounders employed by the authors of the included studies; the quality of control for confounding for each paper would then be scored using a Likert scale to indicate the extent to which study authors controlled for potential confounders. An overview of the confounders for each outcome of interest, along with the sources used to inform the review team, is available in Appendix A.

Quality assessment ratings were not used to exclude studies from the analysis, but instead were used to describe the main strengths and limitations of the studies.

2.6.2 Case-control studies

Case-control studies were assessed for methodological quality by a single reviewer (JL) using the NHLBI's 13-item quality assessment tool for case-control studies [40]. The reviewer extracted verbatim text from each paper to support the reason for each score.

The tool is presented in full in Appendix G. For each study, an overall quality rating was calculated using a bespoke system, based on essential criteria for high-quality case-control studies [39]. The items chosen concerned aspects of studies that were most likely to introduce bias to the results through unrepresentative sampling, quality of matching, measurement of exposure, and confounding. Five items from the tool were selected and scored as outlined in Table 5. The same thresholds for low (<3.0), moderate (3.0), and high (≥3.5) quality were set as for the ecological, cross-sectional, and cohort studies (see Section 2.6.1). Quality assessment ratings were verified independently by a second reviewer (KL). We treated confounders in the manner described for ecological, cross-sectional, and cohort studies (see Section 2.6.1). Quality assessment ratings were not used to exclude studies from the analysis, but instead were used to describe the main strengths and limitations of the studies.

Table 5 Overall quality rating calculation for case-control studies

Item	Scoring
4. Did the authors include a sample size justification?	Yes: 1.0 No: 0.0
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes: 1.0 Partly: 0.5 No: 0.0
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes: 1.0 No: 0.0
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	Yes: 1.0 Partly: 0.5 No: 0.0
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Almost all: 1.0 Partial: 0.5 Limited: 0.0 None: 0.0

Note: Responses of "Not reported" and "Not applicable" were scored 0.0 for each item. For item 13, key potential confounding variables were identified based on established risk factors for the condition under consideration (see Appendix A); while some studies controlled for a large number of variables in their models, only these key confounding variables were considered for item 13.

2.7 Data synthesis

2.7.1 Quantitative synthesis

For each outcome of interest, we carried out an assessment of the feasibility of meta-analysis following published guidance [42,43]. Studies were grouped first by outcome, then by study type. Following this, for each group of studies, comparability on the following variables was assessed in order:

1. Outcome measures (based on definition and methods of measurement)
2. Populations (based on inspection of inclusion criteria and baseline participant characteristics)
3. Exposures (based on fluoride dose and duration of exposure), and
4. Comparators (based on fluoride dose and duration of exposure/non-exposure).

2.7.2 Narrative synthesis

For each outcome of interest, data were narratively summarised and findings for each outcome or measure for each study were tabulated. We have described the study designs (including suitability for assessing causality), the characteristics of each study, the quality of the study (including limitations leading to bias), the study findings for each outcome, and a concluding statement on the evidence for each outcome.

2.7.3 Applying the Grading of Recommendations, Assessment, Development and Evaluations approach to outcomes

The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) system [44] was employed in order to grade the quality of evidence and strength of the recommendations. While the quality assessment process described in Section 2.6 rates the quality of individual studies, the GRADE approach is used to rate the quality of evidence for each outcome across the studies (i.e. for a body of evidence related to a given outcome, not an individual study).

Under the GRADE system, the initial certainty of the evidence is determined based on study design, with randomised controlled trials providing a high degree of certainty and observational studies providing a low degree of certainty. The level of certainty is then adjusted upwards or downwards based on a number of factors. Ultimately, a body of evidence related to an outcome receives one of four grades: high, moderate, low, or very low, reflecting the level of confidence we may have that the true effect is similar to, or substantially different from, the estimate of the effect.

Following the GRADE approach, we downgraded the quality of the evidence considering five criteria (risk of bias, inconsistency, indirectness, imprecision, and publication bias), and for outcomes where the five criteria were met, we upgraded the quality of the evidence based on three criteria (large effect, dose response, and opposing bias and confounders). Each study starts at 10 points and can lose 0, 1, or 2 points for each of the five downgrading criteria. However, if all five criteria are met, it can gain an additional 1 or 2 points for large effect, and 1 point for dose response and/or opposing bias.

The reasons for downgrading are:

1. Risk of bias, which takes account of study design considering the hierarchy of evidence and the methodological quality of the study
2. Inconsistency, which considers both clinical and statistical heterogeneity that cannot be controlled for in the analysis
3. Indirectness, which considers the comparator intervention and whether it is the current gold standard or it is being used as a proxy; indirectness also considers the population, intervention, and outcome
4. Imprecision, which takes account of the size of the variance and the optimal effect size and is closely related to sample size and the number of events of interest, and
5. Publication bias, which is a systematic underestimation or overestimation of the underlying beneficial or harmful effect due to the selective publication of studies.

The decision to upgrade should only rarely be made if serious limitations are present in any of these areas and should only be made after full consideration – and in the context – of reasons to downgrade. The reasons for upgrading are:

1. Large or very large estimates of the magnitude of an intervention or exposure effect

2. The presence of a dose–response gradient, which may increase confidence in the findings of observational studies, and
3. Where all plausible residual confounding from observational studies may be working to increase or decrease the demonstrated effect, if no effect was observed.

3 Findings

3.1 Search results

A total of 30 studies reported across 37 papers meeting the review eligibility criteria were identified from all stages of the search process (see the overview of literature search results in Appendix I; see the full list of included papers in Appendix J; see Figure 4 for PRISMA flow diagram). Two studies were reported by Lehmann *et al.* in their 1998 paper [31]: a cross-sectional study of bone mineral density and an ecological study of hip fractures. These have been counted as two separate studies due to their distinct designs.

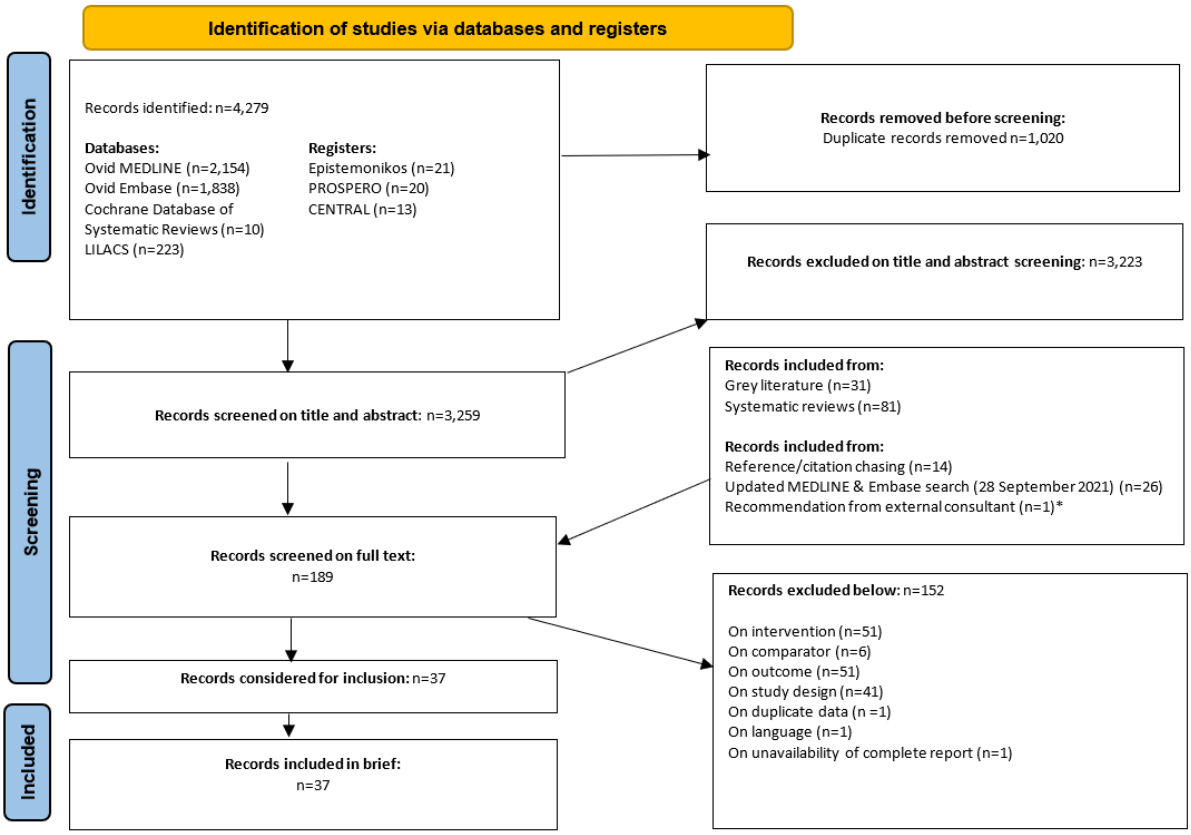

Figure 4 PRISMA flow diagram

*See Section 2.4.2.7

Source: Page *et al.*, 2021 [28]

3.2 Overview of included studies

The countries of origin for the 37 papers and 30 studies included are listed in Table 6. The majority of studies (n=16, 53%) were completed in North America, and 10 (33%) were completed in Europe. Details of the CWF schemes in place in each country, both presently and historically, are provided in Appendix K.

Table 6 Study locations

Country	Number of papers
Canada	12 papers from 6 studies
Finland	1 paper from 1 study
Germany	1 paper with 2 studies
Ireland	2 papers from 2 studies
New Zealand	3 papers from 3 studies
South Korea	1 paper from 1 study
Spain	1 paper from 1 study
United Kingdom	4 papers from 4 studies
USA	12 papers from 10 studies
Total	**37 papers from 30 studies**

A number of studies or papers are linked, either because they analysed data from the same dataset or because they expand on or update the same earlier work. These linked papers are highlighted wherever they arise in the sections reporting on fluoride and its associations with systemic health; it is important for the reader to be aware of these links when interpreting the aggregated evidence, so that the same data are not counted twice.

The following is a summary of the linked papers included in the review:

- The study by Cauley *et al.* (1995) [45] on bone mineral density (BMD) was updated with an increased sample size in Phipps *et al.* (2000) [46] (Study of Osteoporotic Fractures Pittsburgh). These studies are designated "Pittsburgh" with a black spot ● in the tables throughout the document.
- The study by Bassin *et al.* (2006) [47] on osteosarcoma was updated with an increased sample size in Kim *et al.* (2020) [48] (Harvard Fluoride Osteosarcoma Study). These studies are designated "Harvard" with a blue spot ● in the tables throughout the document.
- Barberio *et al.* (2017a) [49] and Riddell *et al.* (2019) [50] analysed data on neurodevelopmental disorders from Statistics Canada's Canadian Health Measures Survey dataset. These studies are designated "CHMS" with a green spot ● in the tables throughout the document.
- Barberio *et al.* (2017b) [51], Cunningham *et al.* (2021) [52], and Malin *et al.* (2018) [53] also analysed data from Statistics Canada's Canadian Health Measures Survey dataset. Cunningham *et al.* analysed data on sleep, while Barberio *et al.* and Malin *et al.* analysed data on thyroid-stimulating hormone (TSH) diagnosis of thyroid conditions. These studies are designated "CHMS" with a green spot ● in the tables throughout the document.
- Three papers (Green *et al.* (2019) [54], Till *et al.* (2020) [55], and Farmus *et al.* (2021) [56]) analysed data on IQ outcomes from 601 of the 2001 mother–child pairs enrolled in the Maternal-Infant Research on Environmental Chemicals longitudinal cohort study. These studies are designated "MIREC" with a red spot ● in the tables throughout the document.

Table 7 displays a summary of the 37 included papers, including country, study design, and outcomes examined. The name of the study is noted in cases where data from the same study are analysed in more than one paper.

Table 7 Summary of included papers

Paper	Country	Study design	Bone health	Neuropsychological outcomes	Cancer	Endocrine conditions	Renal conditions	Birth or birthing abnormalities	Infant abnormalities	All-cause mortality
Arnold *et al.* (1997) [57]	Canada	Ecological or correlational study	✓							
Barberio *et al.* (2017a) [49] (CHMS) 🟢	Canada	Cross-sectional survey		✓						
Barberio *et al.* (2017b) [51] (CHMS) 🟢	Canada	Cross-sectional survey				✓				
Bassin *et al.* (2006) [47] (Harvard) 🔵	USA	Case-control study			✓					
Blakey *et al.* (2014) [58]	Great Britain (England, Wales, and Scotland)	Ecological or correlational study			✓					
Broadbent *et al.* (2015) [59]	New Zealand	Prospective cohort study		✓						
Cauley *et al.* (1995) [45] (Pittsburgh) ⚫	USA	Cross-sectional survey	✓							
Chachra *et al.* (2010) [60]	Canada	Ecological or correlational study	✓							
Cohn (1992) [61]	USA	Ecological or correlational study			✓					
Comber *et al.* (2011) [20]	Ireland	Ecological or correlational study			✓					

Paper	Country	Study design	Bone health	Neuropsychological outcomes	Cancer	Endocrine conditions	Renal conditions	Birth or birthing abnormalities	Infant abnormalities	All-cause mortality
Crnosija *et al.* (2019) [62]	USA	Ecological or correlational study			✓					
Cunningham *et al.* (2021) [52] (CHMS) 🟢	Canada	Cross-sectional survey				✓				
Danielson *et al.* (1992) [63]	USA	Ecological or correlational study	✓							
Dick *et al.* (1999) [64]	New Zealand	Case-control study with ecological assignment of CWF status							✓	
Farmus *et al.* (2021) [56] (MIREC) 🔴	Canada	Prospective cohort study		✓						
Green *et al.* (2019) [54] (MIREC) 🔴	Canada	Prospective cohort study		✓						
Hrudey *et al.* (1990) [65]	Canada	Ecological or correlational study			✓					
Ibarluzea *et al.* (2021) [38]	Spain	Prospective cohort study		✓						
Jacobsen *et al.* (1992) [66]	USA	Ecological or correlational study	✓							
Jacobsen *et al.* (1993) [67]	USA	Ecological or correlational study	✓							
Kim *et al.* (2020) [48] (Harvard) 🔵	USA	Case-control study			✓					

Paper	Country	Study design	Bone health	Neuropsychological outcomes	Cancer	Endocrine conditions	Renal conditions	Birth or birthing abnormalities	Infant abnormalities	All-cause mortality
Kröger *et al.* (1994) [32]	Finland	Cross-sectional survey	✓							
Lee *et al.* (2020) [68]	South Korea	Ecological or correlational study	✓		✓					
Lehmann *et al.* (1998)* [31]	Germany	Cross-sectional survey (bone mineral density)	✓							
		Ecological or correlational study (fracture)	✓							
Lowry et al. (2003) [69]	England	Ecological or correlational study						✓		
Mahoney et al. (1991) [70]	USA	Ecological or correlational study			✓					
Malin *et al.* (2018) [53] (CHMS) 🟢	Canada	Cross-sectional survey				✓				
McGuire *et al.* (1991) [71]	USA	Matched case-control study			✓					
National Fluoridation Information Service (2013) [72]	New Zealand	Ecological or correlational study			✓					

Paper	Country	Study design	Bone health	Neuropsychological outcomes	Cancer	Endocrine conditions	Renal conditions	Birth or birthing abnormalities	Infant abnormalities	All-cause mortality
O'Sullivan and O'Connell (2014) [73]	Ireland	Cross-sectional survey; water fluoridation status assigned on ecological/ population basis	✓							
Peckham *et al.* (2015) [11]	England	Ecological or correlational study				✓				
Phipps *et al.* (2000) [46] (Pittsburgh) ●	USA	Cross-sectional survey	✓							
Riddell *et al.* (2019) [50] (CHMS) ●	Canada	Cross-sectional survey		✓						
Suarez-Almazor *et al.* (1993) [30]	Canada	Ecological or correlational study	✓							
Till *et al.* (2020) [55] (MIREC) ●	Canada	Prospective cohort study		✓						
Young *et al.* (2015) [41]	England	Ecological or correlational study	✓		✓		✓	✓		✓
Zhang *et al.* (2019) [74]	USA	Ecological or correlational study						✓		
Total			**14**	**7**	**12**	**4**	**1**	**3**	**1**	**1**

* Two studies were reported by Lehmann *et al.* in their 1998 paper [31]: a cross-sectional study of bone mineral density and an ecological study of hip fractures. These have been counted as two separate studies due to their distinct designs.

Many papers were based on ecological studies (19), cross-sectional surveys (10), or case-control studies (4). The Glossary of terms presents a description of the study designs utilised in these papers, as this is important for determining the level of evidence each study design provides and its possible contribution to causality.

One (O'Sullivan and O'Connell (2014) [73]) of the 10 cross-sectional surveys and one (Dick *et al.* (1999) [64]) of the 4 case-control studies allocated exposure on an ecological or population basis (i.e. assigning fluoridation status to individuals according to the prevalence of CWF in their regions, rather than the known CWF status of their individual addresses or address history), indicating that they adhered to an ecological rather than cross-sectional survey design. This is important, as ecological studies are vulnerable to 'ecological fallacy', whereby it is not known whether the individuals who were exposed to artificially fluoridated water, in this case, were the same individuals who experienced the outcome (or disease) of interest. The findings of ecological studies are used to develop theories rather than to test them. Notably, ecological studies do not usually appear in the hierarchy of evidence.

The remaining nine cross-sectional surveys and three case-control studies are based on exposure and outcome data collected from individuals and may be considered true cross-sectional surveys and case-control studies. Cross-sectional surveys studies establish a definitive temporal sequence, as data on both exposure (in this case, artificially fluoridated water) and outcome are collected at the same time and it is not known which came first – the exposure or outcome. Case-control studies are useful for investigating rare diseases such as osteosarcoma. Of note is the fact that case-control studies and retrospective cohort studies suffer from recall bias, which is a problem for all the included case-control studies, but it is more severe in the case of Bassin *et al.*, due to the long delay between diagnosis and the data collection on exposure.

The sections covering fluoride and its associations with systemic health explore the evidence on the influence of artificially fluoridated water on eight outcomes of interest:

1. Bone health
2. Neuropsychological outcomes
3. Cancer
4. Endocrine conditions
5. Renal conditions
6. Birth or birthing abnormalities
7. Infant abnormalities, and
8. All-cause mortality.

Appendix A presents the theoretical basis for a link to fluoride for each of these outcomes, along with the factors expected to be controlled for in predictive models.

The results of the feasibility assessment (see Appendix H) indicated that meta-analysis was not possible for any outcomes. The reasons included differing study designs, the number of studies measuring the same outcome, limitations of the included studies, and differing outcomes of studies and their various means of measurement. The following sections therefore present a narrative synthesis of the findings for each outcome.

3.3 Bone health

Concerns about fluoride's effects on the skeletal system focus on bone mass, bone mineral density, skeletal fluorosis, and bone fracture. Fluoride is readily incorporated into the crystalline structure of bone and accumulates over time. Bone mineral density (BMD) is a measure of the amount of minerals (mostly calcium and phosphorous) contained in a certain volume of bone. Fluoride at high levels (e.g. therapeutic doses for treatment of osteoporosis ≥20 to 60 mg/day) increases bone density [75], but may reduce bone strength [76] and appears to exacerbate the growth of osteophytes present in the bones and joints, resulting in joint stiffness and pain. In severe cases it progresses, causing skeletal fluorosis, which is a bone and joint condition associated with prolonged exposure to high concentrations of fluoride. Skeletal fluorosis is typically seen in regions (such as parts of India, China, and Tanzania) with high levels of natural fluoride (>3 ppm) in groundwater [25]. These countries experience fluoride levels that can be up to 12 times higher than the level in CWF schemes. In the late 20th century, some researchers thought that artificially fluoridated water might be protective against bone fracture, whereas other researchers thought that fluoride exposure may increase the risk of hip fracture in persons treated for osteoporosis [31].

Thirteen studies reported across 13 papers [30–32,41,45,46,57,60,63,66–68,73] examined the associations between fluoridated water and bone health. Bone characteristics (including BMD and osteoporosis) were the outcome of interest in 8 papers, while incidence of fractures (most commonly hip fractures) was examined by 10 papers.

3.3.1 Bone characteristics

Eight papers [31,32,45,46,57,60,68,73], describing seven studies, examined the associations between fluoridated water and bone characteristics. BMD was the outcome of interest in seven papers (six studies), while osteoporosis was examined by one paper (one study).

3.3.1.1 Study characteristics

The summary characteristics of the eight papers that examined bone characteristics are presented in Table 8 (see Appendix L for full study characteristics). Four papers presented data from three cross-sectional surveys, three papers were ecological studies, and one paper was a cross-sectional study that allocated exposure on an ecological or population basis, which should therefore be treated as an ecological study for the purposes of contributing evidence for causality.

Four of the eight papers [31,57,60,73] examined differences in BMD between fluoridated and non-fluoridated areas, three papers [32,45,46] based on two studies examined differences in BMD between groups of participants according to length of exposure to fluoridated water, and one paper [68] examined differences in the incidence of osteoporosis according to the extent of implementation of fluoridation.

None of the studies examined skeletal fluorosis; although this is an important outcome related to fluoride concentration in water, as outlined in Section 1.1.1, it tends to occur only in areas with natural fluoride concentrations above 3 ppm. These areas have been excluded from consideration in this review.

It is important to note that the 2000 paper by Phipps *et al.* [46] presents an update of the study by Cauley *et al.* published in 1995 [45] (Study of Osteoporotic Fractures Pittsburgh). The original and updated phases of the Pittsburgh study were carried out by overlapping research teams, and the 2000 paper presents data from a larger sample over a longer period, with the intention of increasing statistical power.

Table 8 Summary of study characteristics for studies examining bone characteristics

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Bone mineral density						
Kröger *et al.* (1994)	Cross-sectional survey Finland	Kuopio Osteoporosis Study: all perimenopausal women aged 47–56 years residing in Kuopio Province, Finland in February 1989; random stratified sample of those willing to undergo bone densitometry	Women using fluoridated drinking water for more than 10 years (1.0–1.2 mg/L)	Women who did not have access to artificially fluoridated drinking water or who had used it for less than 10 years. The fluoride content of drinking water is low (0.0–0.3 mg/L).	BMD of spine and neck of femur	Fluoridated: 969; non-fluoridated: 2,253
Cauley *et al.* (1995) (Pittsburgh) ●	Cross-sectional survey USA	Women aged 65 years or over, excluding black women (due to reduced incidence of hip fractures) and women unable to walk without the assistance of another person or who had bilateral hip replacements; most recruited from voter registration lists for ZIP codes within a 25-mile radius of Monessen, Pennsylvania, USA	Years of exposure to fluoridated community water supplies recorded for each participant; exposure duration range: 1–38 years; mean fluoride concentration 1.01 ppm (±0.21 SD) for fluoridated public water	Zero years of exposure to fluoridated community water supplies recorded for each participant	Bone mineral content and density for the spine and hip and at the midpoint and ultradistal radius and calcaneus. Spinal and non-spinal fractures were also recorded.	Zero years of fluoride exposure: 1,248; 1–10 years of fluoride exposure: 438; 11–20 years of fluoride exposure: 198; and >20 years of fluoride exposure: 192
Arnold *et al.* (1997)	Ecological or correlational study Canada	Females aged 18–25 years. All subjects had not travelled outside of their resident city in Canada for more than 4 years. Individuals with bone-affecting disorders, use of potential bone-affecting medications, long-term use of fluoride supplements, a history of amenorrhoea (fewer than three	Saskatoon, Saskatchewan, Canada, which has had supplemental fluoride in its water since 1954, at a level of approximately 1.0 mg/L; duration of exposure: >4 years	Regina, Saskatchewan, Canada, which has never had supplemental fluoride in its water supply and has a naturally occurring fluoride level of <0.12 to 0.15 mg/L in its water	BMD for the total body, lumbar spine, and proximal femur	Total: 57 (BMD fluoridated: 33; BMD non-fluoridated: 24)

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		menses per year), and those who were currently pregnant were excluded.				
Lehmann *et al.* (1998)	Cross-sectional survey Germany	BMD analysis: volunteer employees at a local hospital in Germany, excluding those who used drugs affecting calcium metabolism and those with disease known to affect bone metabolism	Chemnitz, Germany, where water was fluoridated to 1 mg/L over a period of 30 years; duration of exposure: >10 years	Halle, Germany, where water was not fluoridated and had a naturally occurring fluoride level of 0.08–0.36 mg/L	BMD	555 (Chemnitz, Germany: 201 women and 41 men, totalling 242; Halle, Germany: 215 women and 98 men, totalling 313) [Authors Table 2]
Phipps *et al.* (2000) (Pittsburgh) ●	Cross-sectional survey USA	Prospective sample of 9,704 white women aged 65 years or over in Portland, Oregon; Minneapolis, Minnesota; Baltimore, Maryland; and the Monongahela Valley, Pennsylvania, USA, recruited from jury selection and voter registration, motor vehicle records, and membership records of health plans. Excluded white women unable to walk without assistance and women who had bilateral hip replacement. Recruitment took place from 1986 to 1988.	Women exposed to fluoridated water continuously for the last 20 years; levels not specified, USA standard target fluoride level was 0.7–1.2 ppm at time of study	Women with no exposure to fluoridated water for the last 20 years; levels not specified	BMD and fractures of the vertebrae, hip, wrist, and humerus	No exposure: 3,218; continuous exposure: 2,563
Chachra *et al.* (2010)	Ecological or correlational study	Patients undergoing total hip arthroplasty in one hospital in each region in Canada between	Fluoridated region (Toronto), 1 ppm; fluoridation in place since	Non-fluoridated region (Montreal), levels not specified	Properties of bone samples: fluoride content	Toronto: 53 (27 female);

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
	Canada	September 1996 and August 2000	1963; duration of exposure not specified		(including density) and structural or mechanical properties of bone	Montreal: 39 (24 female)
O'Sullivan and O'Connell (2014)	Cross-sectional survey; water fluoridation status assigned on ecological/population basis Ireland	Nationally representative sample of people aged 50 years or over (and their spouses and partners of any age) resident in Ireland	Percentage of households with fluoridated water supply in electoral area, specific levels not specified. Each local authority was responsible for fluoridating its own water supply within the range of 0.6–0.8 ppm between 2002 and time of data collection. Prior to 2002, the target range for fluoridation was 0.8–1.0 ppm. Fluoridation status was based on 2006 Census address data and 2006 local government water source data; duration not specified; fluoridation started in 1964.	Fluoride levels in non-fluoridated areas not specified, generally not more than 0.3 ppm in Ireland. According to the 2006 Census, around 84% of households have fluoridated water supplies, which is unsurprising given that all the main urban areas receive local government water supplies.	BMD and body mass index	4,977 people aged 50 years and over
Osteoporosis						
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in	7 areas that did not receive CWF.	Osteoporosis	CWF: 4,406,021; non-CWF: 2,270,959

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
			11 areas, withdrawn in 2004 in all areas			

3.3.1.2 Quality assessment

A summary of the quality assessment of the eight papers that examined bone characteristics, using the NHLBI's quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 9 (see Appendix M for the full quality assessment). Four of the eight papers received a rating of low quality for their conduct, three received a rating of moderate quality, and one received a rating of high quality.

It is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality. In addition, cross-sectional studies are useful for estimating prevalence, planning and evaluating health services, and identifying theoretical relationships, but cannot prove or disprove causality, as they collect exposure and outcome information at the same time.

Table 9 Summary of quality assessment ratings for studies examining bone characteristics

Item	Kröger *et al.* (1994)	Cauley *et al.* (1995) (Pittsburgh) ●	Arnold *et al.* (1997)	Lehmann *et al.* (1998) (BMD)	Phipps *et al.* (2000) (Pittsburgh) ●	Chachra *et al.* (2010)	O'Sullivan and O'Connell (2014)	Lee *et al.* (2020)
3. Was the participation rate of eligible persons at least 50%?	No	No	Not reported	Yes	Yes	Not reported	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	No	No	No	Partly	No	No	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Almost all	Almost all	Almost all	Almost all	Almost all	None	Partial	Limited
Quality rating	3.0 (moderate)	2.0 (low)	2.0 (low)	3.0 (moderate)	3.5 (high)	1.0 (low)	2.5 (low)	3.0 (moderate)

3.3.1.3 Findings

The summary findings of the eight papers (examining seven studies) that examined bone characteristics (including BMD and osteoporosis) are presented in Table 10 (see Appendix N for full findings). Specific statistics are noted in the summary findings tables only for statistically significant analyses. The findings for the influence of fluoridation on BMD were mixed, as outlined below:

- For the lumbar spine, an ecological study [57] found an association between lower BMD and exposure to water fluoridation. The initial analysis of the Pittsburgh cross-sectional study by Cauley *et al.* (1995) [45] found no association; however, the updated version with a larger sample size (Phipps *et al*., 2000) [46] found an association between higher BMD and exposure to water fluoridation, as did one additional cross-sectional study [32]. One further cross-sectional study [31] found no association.
- For the ultradistal radius, one cross-sectional study described in two papers [45,46] found an association between lower BMD and water fluoridation among women aged 65 years or over.
- For the total hip, one cross-sectional study [45] found no association between BMD and water fluoridation status.
- For the femoral head, one ecological study [60] found an association between higher BMD and exposure to water fluoridation.
- For the femoral neck, one cross-sectional study [31] found no association between BMD and water fluoridation status. The initial analysis of the Pittsburgh cross-sectional study by Cauley *et al.* (1995) [45] found no association, while the updated version with a larger sample size (Phipps *et al*., 2000) [46] found an association between higher BMD and exposure to water fluoridation among women aged 65 years or over. One further cross-sectional study [32] also found an association between higher BMD and exposure to water fluoridation among postmenopausal women.
- For Ward's triangle (a radiolucent area between principal compressive, secondary compressive, and primary tensile trabeculae in the neck of the femur), the initial analysis of the Pittsburgh cross-sectional study (Cauley *et al.* (1995)) [45] found no association, while the updated version with a larger sample size (Phipps *et al*., 2000) [46] found an association between higher BMD and exposure to water fluoridation. One further cross-sectional study [31] found a similar association between higher BMD and exposure to water fluoridation.
- For the trochanter of the femur, one cross-sectional study (Lehmann *et al.*, 1998) [31] found no association between BMD and water fluoridation status. The initial analysis of the Pittsburgh cross-sectional study (Cauley *et al.* (1995)) [45] also found no association; however, the updated version of the study with a larger sample size (Phipps *et al*., 2000) [46] found an association between higher BMD and exposure to water fluoridation among women aged 65 years or over.
- For the intertrochanter of the femur, one cross-sectional study described across two papers [45,46] found no association between BMD and water fluoridation status.
- For the femur (area not specified), an ecological study [57] found no association between BMD and water fluoridation status.
- For bone stiffness of the foot, a cross-sectional survey with exposure status assigned on an ecological basis [73] found no association between BMD and water fluoridation status.

Additionally, one ecological study found no association between the incidence of osteoporosis and water fluoridation status [68].

The nature of the measure of exposure (i.e. whether exposure was measured according to residence in fluoridated areas or duration of exposure to fluoridated water) did not appear to influence how likely studies were to find possible relationships.

The nature of the measure of outcome is also worth noting, as a majority of studies used photon absorptiometry, either single-photon or dual-energy X-ray absorptiometry (DXA). Photon absorptiometry was first described in the 1960s. The method is based on the relationship between photon attenuation and tissue density; a very dense tissue, such as bone, absorbs more photons than soft tissue. Advancements in the method were made by adding a second photon source of a different energy, allowing soft tissue and bone tissue to be assessed separately. This also removed the necessity for the site of interest to be submerged in water, allowing for central sites (e.g. the spine and femur) as well as peripheral sites (e.g. the radius) to be assessed. In the 1980s, the concept was adapted to use X-rays, rather than photons, which are similarly attenuated in proportion to tissue density. DXA reduced the radiation exposure involved in assessment. Further advancements have also made scan times significantly faster [77].

In a number of studies comparing measurements of the radius, single-photon absorptiometry and DXA measurements of bone mineral density have been shown to be tightly correlated and provide equivalent diagnostic information; however, DXA measurements have been shown to be more reproducible [78–81].

Of the studies included in this review, approximately half of the comparisons using DXA revealed statistically significant differences; this was true for only a quarter of the comparisons using single photon absorptiometry. However, single photon absorptiometry was used in only the Pittsburgh study, presented in two papers [45,46], while DXA was used in four papers [31,32,46,57], therefore representing a larger sample of individuals. Therefore, it would not be appropriate to interpret these differences as an indication that DXA is a more sensitive measure; as mentioned above, the two instruments have been demonstrated to provide equivalent diagnostic information.

Overall, the evidence from these studies for the influence of CWF on BMD is mixed; although a number of studies found associations between higher BMD and exposure to water fluoridation in certain skeletal areas (such as the lumbar spine), contradictory findings also exist, and a large number of analyses found no association. Therefore, no theoretical relationship has been firmly established. A high-quality prospective longitudinal study based on individual-level exposures, taking account of all potential confounding factors and effect modifiers, is required in order to strengthen the evidence base on BMD.

Table 10 Summary of findings for studies examining bone characteristics

Paper	Comparisons	Method of measurement	Summary of findings
Bone mineral density: Total body			
Arnold *et al.* (1997)	Fluoridated versus (vs.) non-fluoridated regions	Dual-energy X-ray absorptiometry (DXA) in array mode	No difference in total body BMD between women raised in fluoridated and non-fluoridated areas
Bone mineral density: Lumbar spine			
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	DXA	Women with continuous exposure had significantly higher BMD of the lumbar spine (p<0.001)
Bone mineral density: Lumbar spine L1–L4			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of the lumbar spine L1–L4
Bone mineral density: Lumbar spine L2–L4			

Paper	Comparisons	Method of measurement	Summary of findings
Kröger *et al.* (1994)	Women using or not using fluoridated drinking water for more than 10 years; subanalyses of premenopausal and postmenopausal women	DXA	Perimenopausal women using fluoridated drinking water for more than 10 years had significantly higher BMD in the spine than those not using fluoridated water; significant differences for premenopausal women (p=0.002), postmenopausal women (p=0.005), and whole sample (p=0.001)
Lehmann *et al.* (1998)	Men and women in fluoridated vs. non-fluoridated regions	DXA	No significant differences in adjusted BMD between regions for either men or women
Bone mineral density: Estimated volumetric lumbar 3			
Arnold *et al.* (1997)	Fluoridated vs. non-fluoridated regions	Volumetric estimate of BMD (bone mineral content divided by estimated volume)	Women raised in a non-fluoridated area had significantly higher BMD of volumetric lumbar 3 (p<0.05)
Bone mineral density: Anterior-posterior lumbar spine			
Arnold *et al.* (1997)	Fluoridated vs. non-fluoridated regions	DXA in array mode	Women raised in a non-fluoridated area had significantly higher density of anterior-posterior lumbar spine (p<0.05)
Bone mineral density: Radius – distal			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of distal radius
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Single photon absorptiometry	Women with continuous exposure had significantly lower density of the distal radius (p=0.002).
Bone mineral density: Radius – proximal			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	Density of proximal radius was significantly lower among those exposed for 1–10 years than among those with 0 years of exposure (p=0.05).
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Single photon absorptiometry	Women with continuous exposure had significantly lower density of the proximal radius (p<0.001).
Bone mineral density: Radius – calcaneus			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of radius calcaneus
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Single photon absorptiometry	No significant difference in density of the radius calcaneus between women with continuous exposure and those with no exposure
Bone mineral density: Total hip			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of total hip
Bone mineral density: Femoral head			
Chachra *et al.* (2010)	Fluoridated vs. non-fluoridated regions	Micrometry	Significantly higher BMD observed in samples from the fluoridated region (p<0.05)

Paper	Comparisons	Method of measurement	Summary of findings
Bone mineral density: Femoral neck			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of femoral neck
Lehmann *et al.* (1998)	Men and women in fluoridated vs. non-fluoridated regions	DXA	No significant differences in adjusted BMD between regions
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Method not specified	Women with continuous exposure had significantly higher density of the femoral neck (p<0.001)
Kröger *et al.* (1994)	Women using or not using fluoridated drinking water for more than 10 years; subanalyses of premenopausal and postmenopausal women	DXA	Significantly higher density of the femoral neck observed for postmenopausal women (p<0.004)
Bone mineral density: Ward's triangle			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of Ward's triangle
Lehmann *et al.* (1998)	Men and women in fluoridated vs. non-fluoridated regions	DXA	Density of Ward's triangle significantly higher for men in fluoridated area (p=0.002); no association among women
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Method not specified	Women with continuous exposure had significantly higher density of Ward's triangle (p=0.002)
Bone mineral density: Trochanter			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of the trochanter
Lehmann *et al.* (1998)	Men and women in fluoridated vs. non-fluoridated regions	DXA	No significant differences in adjusted BMD between regions
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Method not specified	Women with continuous exposure had significantly higher density of the trochanter (p<0.001)
Bone mineral density: Intertrochanter			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 0 vs. 1–10 vs. 11–20 vs. >20	Single photon absorptiometry	No association between fluoride exposure and density of the intertrochanter
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Method not specified	No significant difference in density of the intertrochanter between women with continuous exposure and those with no exposure
Bone mineral density: Proximal femur			
Arnold *et al.* (1997)	Fluoridated vs. non-fluoridated regions	DXA in array mode	No difference in density of proximal femur between women raised in fluoridated and non-fluoridated areas
Bone stiffness: Non-dominant foot			

Paper	Comparisons	Method of measurement	Summary of findings
O'Sullivan and O'Connell (2014)	All participants; sensitivity analysis conducted for non-fully urbanised, men only, women only, and those aged under 55 years	Quantitative ultrasound	No association between prevalence of households with fluoridated water and probability of normal bone density
Osteoporosis			
Lee *et al.* (2020)	Risk over time in fluoridated and non-fluoridated regions for total sample, men, and women	Osteoporosis incidence data from National Health Insurance Service	Relative risks increased over time but did not increase in CWF area compared with non-CWF area

3.3.2 Fractures

Ten papers [30–32,41,45,46,63,66–68] presented data from 9 studies that examined the association between water fluoridation status and the incidence of a range of fractures, most commonly hip fractures.

3.3.2.1 Study characteristics

A summary of the characteristics of the 10 papers that examined fractures is presented in Table 11 (see Appendix L for full study characteristics). Seven papers were based on ecological studies and three papers were based on two cross-sectional studies.

It is important to note that the 2000 paper by Phipps *et al.* [46] presents an update of the study by Cauley *et al.* published in 1995 [45]. The original and updated phases of the Pittsburgh study were carried out by overlapping research teams, and the 2000 paper presents data from a larger sample over a longer period, with the intention of increasing statistical power.

Six of the 10 papers [30–32,41,63,66] compared fracture incidence between areas with and without fluoridated water, 2 papers [45,46] based on one cross-sectional study examined differences in fracture incidence between groups of participants according to length of exposure to fluoridated water, 1 paper [67] compared differences in fracture incidence before and after the introduction of CWF, and 1 paper [68] examined differences in the incidence of osteoporosis according to the extent of implementation of fluoridation. Several studies excluded incidence of traumatic hip fractures from consideration at either the selection stage or the analysis stage, as these have a different aetiology.

Table 11 Summary of study characteristics for studies examining fractures

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Fractures						
Danielson *et al.* (1992)	Ecological or correlational study USA	Cases of hip fractures in those aged 65 years or over requiring hospitalisation 1984–1990 in rural communities in Utah, USA (records of Medicare admissions and discharges) (excluding those aged under 65 years; surgical revision of hip fracture; cases where fracture is possibly due to metastatic cancer; or represented a second fracture); age-specific populations for those aged 65 years or over drawn from the Utah Peer Review Organisation files of Medicare recipients, obtained from annual census counts carried out by the Social Security Administration	One rural community in Utah, USA, with fluoridated water (1 ppm) since 1966; duration of exposure: 24 years	Two rural communities in Utah, USA, with non-fluoridated water (<0.3 ppm)	Hip fracture	Fluoridated community: 84 (65 females, 77.4%); non-fluoridated community: 162 (130 females, 80.3%)
Jacobsen *et al.* (1992)	Ecological or correlational study USA	Cases of hip fracture recorded by the Health Care Financing Administration and the United States Department of Veterans Affairs for white women and men aged 65 years and over for the period 1984–1987; excluding those aged under 65 years, of non-white race, located in Puerto Rico or missing a ZIP code, cases of second fracture of hip, cases where fracture was secondary to metastatic or primary neoplastic disease, or cases where primary discharge diagnosis was for late effects of hip fracture or orthopaedic aftercare	Counties that were >50% urban; natural fluoride level was <0.3 ppm; <10% of the population was served with fluoridated water prior to change, which increased to >67% of the population served with fluoridated water within a period of 3 years; duration of exposure not specified	Counties that were >50% urban; <10% of the population was served with fluoridated water; natural fluoride levels were <0.3 ppm during 1985	Hip fracture	Not reported
Suarez-Almazor *et al.* (1993)	Ecological or correlational study Canada	Cases of hip fracture (discharge diagnosis) in individuals aged 45 years or over living in Edmonton or Calgary, Alberta, Canada, who were admitted to hospital in Alberta between 1981 and 1987; population	Hip fracture admissions in Edmonton, where water has been fluoridated to 1.0 mg/L since 1967; duration of exposure not specified	Hip fracture admissions in Calgary, where water is not fluoridated, and natural levels are	Hip fracture	Primary hip fractures in Edmonton: 2,479; primary hip fractures in Calgary: 2,392

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		estimates from the Alberta Bureau of Statistics with linear interpolation		on average about 0.3 mg/L		
Jacobsen *et al.* (1993)	Ecological or correlational study USA	All incident hip fractures occurring in Rochester, Minnesota, USA from 1950 to 1969 among persons aged 50 years and over. Data taken from Mayo Clinic master index of all diagnoses and surgical procedures.	Rochester, Minnesota, USA, where water fluoridation was introduced in 1960, with levels maintained at 1.1 ppm; duration of exposure: >10 years	Rochester, Minnesota, USA, prior to introduction of water fluoridation in 1960	Hip fracture	651 (383 women)
Kröger *et al.* (1994)	Cross-sectional survey Finland	Kuopio Osteoporosis Study: all perimenopausal women aged 47–56 years residing in Kuopio Province, Finland in February 1989; random stratified sample of those willing to undergo bone densitometry	Women using fluoridated drinking water for more than 10 years (1.0–1.2 mg/L)	Women who did not have access to artificially fluoridated drinking water or who had used it for less than 10 years. The fluoride content of drinking water is low (0.0–0.3 mg/L).	Incidence of wrist fractures, ankle fractures, other fractures, and all fractures	Fluoride group: 969; non-fluoride group: 2,253
Cauley *et al.* (1995) (Pittsburgh) ●	Cross-sectional survey USA	Women aged 65 years or over, excluding black women (due to reduced incidence of hip fractures) and women unable to walk without the assistance of another person or who had bilateral hip replacements; most recruited from voter registration lists for ZIP codes within 25-mile radius of Monessen, Pennsylvania, USA	Years of exposure to fluoridated community water supplies recorded for each participant; exposure duration range: 1–38 years; mean fluoride concentration 1.01 ppm (±0.21 SD) for fluoridated public water	Zero years of exposure to fluoridated community water supplies recorded for each participant	Bone mineral content and density for the spine and hip and at the midpoint and ultradistal radius and calcaneus. Spinal and non-spinal fractures were also recorded.	Zero years of fluoride exposure: 1,248; 1–10 years of fluoride exposure: 438; 11–20 years of fluoride exposure: 198; and >20 years of fluoride exposure: 192

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Lehmann *et al.* (1998)	Ecological or correlational study Germany	Cases: patients aged 35 years or over with hip fractures admitted to local hospitals from 1987 to 1989, excluding those admitted after trauma or pathological fractures. Population: residents of two communities in Germany – capitals of districts, industrial centres of respective regions, population estimates by 5-year age groups between 1987 and 1989 obtained from each city's Bureau of Statistics.	Chemnitz, Germany, where water was fluoridated to 1 mg/L over a period of 30 years; duration of exposure: >10 years	Halle, Germany, where water was not fluoridated and has a naturally occurring fluoride level of 0.08–0.36 mg/L	Hip fracture	612 in Chemnitz, Germany, and 640 in Halle, Germany
Phipps *et al.* (2000) (Pittsburgh) ●	Cross-sectional survey USA	Prospective sample of 9,704 white women aged 65 years or over in Portland, Oregon; Minneapolis, Minnesota; Baltimore, Maryland; and the Monongahela Valley, Pennsylvania, USA, recruited from jury selection and voter registration, motor vehicle records, and membership records of health plans. Excluded white women unable to walk without assistance and women who had bilateral hip replacement. Recruitment took place from 1986 to 1988.	Women exposed to fluoridated water continuously for the last 20 years; levels not specified, USA standard target fluoride level was 0.7–1.2 ppm at time of study	Women with no exposure to fluoridated water for the last 20 years; levels not specified	BMD and fractures of the vertebrae, hip, wrist, and humerus	No exposure: 3,218; continuous exposure: 2,563
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, aims to fluoridate to 1 ppm; duration of exposure not specified	Non-fluoridated areas, levels not specified	Hip fracture	Population: Fluoridated areas 37,971,918, non-fluoridated areas 274,884,530 Cases: Fluoridated areas 45,219, non-fluoridated areas 303,848

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in 11 areas, withdrawn in 2004 in all areas	7 areas did not receive CWF.	Hip fracture	CWF: 4,406,021; non-CWF: 2,270,959

3.3.2.2 Quality assessment

A summary of the quality assessment of the 10 papers that examined fractures, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 12 (see Appendix M for the full quality assessment). One paper received a rating of low quality, one received a rating of moderate quality, and the remaining eight received a rating of high quality. It is important to note that none of the included studies controlled for osteoporosis, which is the leading risk factor for hip fracture. Seven of the papers were ecological studies, which are used to develop theories rather than test them. The Cauley *et al.* paper [45] and the Phipps *et al.* paper [46], which was an update on the study in the Cauley *et al.* paper, present data from a cross-sectional survey, which cannot establish a definitive temporal sequence, as data on both exposure and outcome are collected at the same time point and it is not known which came first – the exposure or outcome.

Table 12 Summary of quality assessment ratings for studies examining fractures

Item	Danielson *et al.* (1992)	Jacobsen *et al.* (1992)	Suarez-Almazor *et al.* (1993)	Jacobsen *et al.* (1993)	Kröger *et al.* (1994)	Cauley *et al.* (1995) (Pittsburgh) ●	Lehmann *et al.* (1998) (Fractures)	Phipps *et al.* (2000) (Pittsburgh) ●	Young *et al.* (2015)	Lee *et al.* (2020)
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes	Yes	Yes	No	Yes	Partly	Yes	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Partial	Partial	Partial	Partial	Almost all	Almost all	Partial	Almost all	Partial	Partial
Quality rating	3.5 (high)	3.5 (high)	3.5 (high)	3.5 (high)	3.0 (moderate)	2.0 (low)	3.5 (high)	3.5 (high)	3.5 (high)	3.5 (high)

3.3.2.3 Findings

The summary findings of the 10 papers that examined fractures are presented in Table 13 (see Appendix N for full findings). There were no associations identified between water fluoridation status and a range of fracture types, including all fractures, osteoporotic fractures, spinal/vertebral fractures, non-spinal fractures, wrist fractures, ankle fractures, non-wrist/non-ankle fractures, and humerus fractures.

Four ecological study papers [30,41,67,68] and one cross-sectional survey paper [45] found no association between water fluoridation and the incidence of hip fracture. Two studies found an association between lower incidence of hip fracture and exposure to water fluoridation, suggesting a possible protective effect; one was a cross-sectional study [46] of women aged 65 years or over and an update of Cauley *et al.*, and the other was an ecological study [31], which only found this association for women aged 85 years or over. Two ecological study papers [63,66] found an higher incidence of hip fracture associated with water fluoridation, suggesting a potential harmful effect; however, the sample sizes and differences in incidences were small in both studies. Overall, the findings were mixed regarding the association between water fluoridation and the incidence of hip fracture, although one-half of the studies found no association between the incidence of hip fracture and water fluoridation status.

Overall, the evidence from these papers for an association between CWF and fracture incidence is mixed, with most analyses pointing to a neutral or, in a few analyses, possible protective effect of fluoridation, although only hip fracture has been extensively studied. It is important to note that none of the included analyses controlled for osteoporosis, which is the leading risk factor for hip fracture. A high-quality prospective longitudinal study based on individual-level exposures, taking account of all potential confounding factors (including osteoporosis) and effect modifiers, is required in order to strengthen the evidence base on fractures.

Table 13 Summary of findings for studies examining fractures

Paper	Comparisons	Method of measurement	Summary of findings
All fractures			
Kröger *et al.* (1994)	Fluoridated vs. non-fluoridated regions	Self-reported fracture since the age of 15 years	No difference in fracture incidence between fluoride and non-fluoride groups
Osteoporotic fracture			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 1–10 vs. 11–20 vs. >20	Non-spine fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	Women exposed for >20 years had about a 25% lower osteoporotic fracture risk, but confidence intervals (Cis) were wide and included 1.0.
Incidental spinal/vertebral fracture			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 1–10 vs. 11–20 vs. >20	Defined as 20% reduction in the vertebral height of the anterior, middle, or posterior dimension of a vertebral body and at least a 4 mm decrease in the vertebral height of a dimension, detected by repeat lateral and lumbar and thoracic vertebral film	No association between fluoride exposure and risk of incidence of vertebral fracture
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Spine fractures, detected by lateral radiographs of thoracic and lumbar spine	Women with continuous exposure to fluoride had a 27% lower risk of vertebral fracture compared with those with no fluoride exposure.

Paper	Comparisons	Method of measurement	Summary of findings
Non-spine/non-vertebral fracture			
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 1–10 vs. 11–20 vs. >20	Non-spine fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	Women exposed for >20 years had about a 25% lower risk of non-spinal fracture, but CIs were wide and included 1.0.
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Non-spine fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	No significant difference between women with continuous exposure compared with those with no fluoride exposure.
Hip fracture			
Danielson *et al.* (1992)	Fluoridated vs. non-fluoridated regions for men and women	Incidence of discharge for hip fracture among those aged 65 years or over	Small but statistically significant higher incidence of hip fractures in the fluoridated area for both men (risk ratio: 1.27 (1.08–1.46)) and women (risk ratio: 1.41 (1.00–1.81))
Jacobsen *et al.* (1992)	Fluoridated vs. non-fluoridated regions for men and women	Incidence of discharge for hip fracture among the white population aged 65 years or over	Small positive ecological association between fluoridation of public water supplies and incidence of hip fracture among the population aged 65 years or over (risk ratios: 1.08 (1.06–1.11) for women; 1.17 (1.13–1.22) for men), although the authors cite numerous methodological concerns regarding accurate measurement and interpret their findings very cautiously.
Suarez-Almazor *et al.* (1993)	Fluoridated vs. non-fluoridated regions for men and women aged 45–64 years, 65 years and over, and total sample	Cases of hip fracture (discharge diagnosis) in individuals aged 45 years or over	Generally no differences observed between the two regions – small difference for men total (rate ratio: 1.12 (1.01–1.24)) and men aged 65 years or over (rate ratio: 1.13 (1.00–1.27)), but the authors judge this unlikely to be meaningful.
Jacobsen *et al.* (1993)	Pre- vs. post-fluoridation for men, women, and whole sample	Incidence of hip fractures (i.e. proximal femur fracture) via Rochester Epidemiology Project and Mayo Clinic master index	No change in risk of hip fracture associated with CWF
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 1–10 vs. 11–20 vs. >20	Hip fractures, self-reported and confirmed by review of copies of radiographs, excluding fractures due to major trauma (e.g. motor vehicle accident)	The relative risk of hip fracture tended to decrease with increasing duration of exposure to fluoride, but the CIs were wide and none of the relative risks were statistically significant.
Lehmann *et al.* (1998)	Fluoridated vs. non-fluoridated regions for men and women	Incidence of hip fracture admissions to local hospitals	No difference in fracture incidence for those aged 35–59 years. No difference in fracture incidence between fluoridated and non-fluoridated areas for those aged 65 years or over, except for women aged over 85 years, for whom fracture incidence was significantly

Paper	Comparisons	Method of measurement	Summary of findings
	across five age groups		lower in fluoridated areas (odds ratio: 1.41 (1.10–1.81), *p*=0.006).
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Hip fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	Women with continuous exposure to fluoride had a 31% lower risk of hip fracture compared with those with no fluoride exposure (relative risk: 0.69 (0.50–0.96), *p*=0.028).
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	Number of hip fracture inpatient consultant episodes per lower super output area level in England recorded in hospital episode statistics between April 2007 and March 2013	No difference in hip fracture incidence between fluoridated and non-fluoridated areas
Lee *et al.* (2020)	Risk over time in fluoridated and non-fluoridated regions for total sample, men, and women	Incidence of hip fracture, data gathered from National Health Insurance Scheme	Relative risks increased over time but did not increase in CWF area compared to non-CWF areas
Wrist fracture			
Kröger *et al.* (1994)	Fluoridated vs. non-fluoridated regions	Wrist fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	No difference in fracture incidence between fluoride and non-fluoride groups
Cauley *et al.* (1995) (Pittsburgh) ●	Years of fluoride exposure: 1–10 vs. 11–20 vs. >20	Wrist fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	No association between fluoride exposure and risk of wrist fracture
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Wrist fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	No significant difference between women with continuous exposure compared with those with no fluoride exposure
Ankle fracture			
Kröger *et al.* (1994)	Fluoridated vs. non-fluoridated regions	Self-reported ankle fracture since the age of 15 years	No difference in fracture incidence between fluoride and non-fluoride groups
Non-wrist/non-ankle fracture			
Kröger *et al.* (1994)	Fluoridated vs. non-fluoridated regions	Self-reported non-wrist/non-ankle fracture since the age of 15 years	No difference in fracture incidence between fluoride and non-fluoride groups
Humerus fracture			
Phipps *et al.* (2000) (Pittsburgh) ●	No exposure vs. continuous exposure for 20 years	Humerus fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)	No significant difference between women with continuous exposure compared with those with no fluoride exposure

3.4 Neuropsychological outcomes

Fluoride is known to cross the placenta and it has been described as a neurotoxin [82]. Concerns have been raised for a long time about its developmental effects on children (see Appendix A). In recent years, concern has been expressed about a possible association between fluoride in drinking water and low IQ levels. This concern has arisen largely due to a group of studies conducted in Asia (in countries such as China, India, and Iran) and Latin America (in Mexico), where fluoride is naturally present in the groundwater at very high levels in certain areas [25]. These studies compared the IQ of children in areas with very high levels of naturally occurring fluoride in the drinking water with the IQ of children in areas of the same country or region with low levels of naturally occurring fluoride in the drinking water. The fluoride levels in the low-fluoride comparison group are generally the same as, or slightly higher than, those found in countries with CWF. Therefore, studies that report findings indicating a lower IQ in children who are drinking fluoridated water are reaching this conclusion from the outcomes of analysis of children drinking water with very high levels of fluoride, compared with children who are drinking water containing fluoride at low levels (similar to CWF levels) [83,84].

Another issue with studies from these countries is that, since fluoride is naturally occurring in these regions, food may also be contaminated with high levels of fluoride from the soil and from coal used to cook the food. The studies are generally of low quality, in that they do not take full account of other factors (also called confounders) that could also cause reduced IQ, such as nutritional status, socioeconomic status, iodine deficiency, and the presence of other chemicals in the groundwater (such as arsenic or lead). Apart from the levels of fluoride in the water, these countries are very different from Ireland with respect to climate, nutritional status, and socioeconomic status. Thus, their findings are not applicable to Ireland or other countries with CWF schemes. For a fuller discussion of these studies and their findings, see the original 2015 HRB review [25].

Seven papers [38,49,50,54–56,59] based on four studies examined the association between fluoridated water and neuropsychological outcomes. Four papers [54–56,59] based on two studies examined IQ in childhood and adulthood, and one additional paper examined aspects of neuropsychological development in infancy and childhood, which conceptually maps closely to IQ [38]. The remaining two papers [49,50], based on one study, examined a cluster of outcomes incorporating attention deficit hyperactivity disorder (ADHD) and its symptoms, and specific learning disabilities (e.g. dyslexia).

A number of points regarding terminology are important here. Although ADHD and dyslexia are often grouped together in this way, ADHD is not itself a specific learning disability, and the neurodevelopmental nature of some specific learning disabilities, including dyslexia, is not universally accepted [85]. In addition, one study examines attention deficit disorder as an outcome of interest; this is an older term for a diagnosis now described as a subtype of ADHD (ADHD – inattentive type). However, in the interest of brevity, we have grouped these outcomes under the heading ‘neurodevelopmental disorders’. We have also preserved the language used by the study authors at the time of writing, in order to provide an accurate description of their findings; therefore, the study authors’ term ‘learning disability’ has been used, rather than ‘specific learning disability’. Similarly, although the authors of the study that examines neuropsychological development do not refer to this construct as IQ, the two are conceptually similar enough that we have grouped them together in our narrative synthesis for ease of comprehension.

A brief primer on the structure of IQ scores may also be useful here. An IQ score or Full Scale IQ (FSIQ) score refers to the global score of cognitive ability on an IQ test. Index scores on IQ tests refer to scores on specific domains of ability, e.g. a test’s verbal index (sometimes called verbal IQ) or numeric index (numeric IQ). Index scores are correlated with FSIQ scores but also have some degree of independence

from them and from other index scores. Performance IQ is an index measuring a range of non-verbal skills, including fluid reasoning, spatial processing, attentiveness to details, and visual-motor integration.

3.4.1 Study characteristics

A summary of the characteristics of the seven papers that examined neuropsychological outcomes is presented in Table 14 (see Appendix L for full study characteristics).

Study designs were varied. The seven papers were based on four studies: three prospective cohort studies and one cross-sectional study. Some papers measured fluoride exposure as both continuous variables (e.g. fluoride concentration in tap water samples or urinary samples) and as dichotomous variables (e.g. residence in a city with or without CWF).

It is important to note that three papers on IQ (Green *et al.* (2019) [54], Till *et al.* (2020) [55], and Farmus *et al.* (2021) [56]) analyse data from the same programme of research; all examine data from 601 of the 2,001 mother–child pairs enrolled in the Maternal-Infant Research on Environmental Chemicals (MIREC) longitudinal cohort study. It should also be noted that these studies examine fluoride intake from tea and coffee as well as from tap water. While the eligibility criteria for this review require that fluoride from tap water be examined in isolation, tea and coffee intake is unlikely to systematically differ across fluoridated and non-fluoridated areas, and so we have chosen not to exclude these studies on this basis.

Additionally, both papers on neurodevelopmental disorders (Barberio *et al.* (2017a) [49] and Riddell *et al.* (2019) [50] analyse data from Statistics Canada's Canadian Health Measures Survey (CHMS) dataset, Cycles 2 (2009–2011) and 3 (2012–2013).

Table 14 Summary of study characteristics for studies examining neuropsychological outcomes

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
IQ/neuropsychological development						
Broadbent *et al.* (2015)	Prospective cohort study New Zealand	Representative sample of children born between 1 April 1972 and 31 March 1973 in Dunedin, New Zealand	Residence in area with CWF (0.7–1.0 ppm), use of 0.5 mg fluoride tablets (ever/never), and/or use of fluoridated toothpaste (always/sometimes/never /unknown) by the age of 5 years; duration of exposure: up to 5 years (preschool years only evaluated)	Residence in area without CWF (0.0–0.3 ppm), use of 0.5 mg fluoride tablets (ever/never), and/or use of fluoridated toothpaste (always/sometimes/never /unknown) by the age of 5 years	IQ	992 (childhood IQ); 942 (adult IQ)
Green *et al.* (2019) (MIREC) 🔴	Prospective cohort study Canada	Pregnant women from 10 cities in Canada, who could communicate in English/French, aged over 18 years, within the first 14 weeks of pregnancy; excluded if there was a known foetal abnormality, medical complications, or illicit drug use during pregnancy; subset of children recruited from 6 cities	Fluoridated water (0.59 mg/L (±0.08 SD)); duration of exposure: 9 months (prenatally) Maternal urinary fluoride (MUF) concentration in fluoridated regions averaged across all three trimesters, adjusted for specific gravity 0.69 mg/L (±0.42 SD)	Non-fluoridated water (0.13 mg/L (±0.06 SD)) MUF concentration in non-fluoridated regions averaged across all three trimesters, adjusted for specific gravity 0.40 mg/L (±0.27 SD)	IQ	512 mother–child pairs with urinary fluoride, IQ, and complete covariate data; 400 mother–child pairs with fluoride intake, IQ, and complete covariate data (non-fluoridated: 238; fluoridated: 162)
Till *et al.* (2020) (MIREC) 🔴	Prospective cohort study Canada	2,001 pregnant women from 10 Canadian cities who could communicate in English or French, were aged over 17 years, and were <14 weeks' gestation; excluded if there was a known foetal	Fluoridated water (0.58 mg/L (±0.08 SD) (breastfed group) or 0.59 mg/L (±0.07 SD) (formula-fed group)); duration of exposure: 30–48 months postnatally	0.13 mg/L (±0.06 SD) (breastfed group) or 0.13 mg/L (±0.05 SD) (formula-fed group)	IQ	Total: 398; breastfed: 200 (fluoridated: 83; non-fluoridated: 117); formula-fed: 198 (fluoridated: 68; non-fluoridated: 130)

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		abnormality or illicit drug use in pregnancy, or if they had any medical complications; 610 children recruited to participate in the developmental follow-up phase of the study from 6 of the cities in the original cohort				
Farmus *et al.* (2021) (MIREC) 🔴	Prospective cohort study Canada	MIREC longitudinal cohort: women in Canadian cities aged 18 years or over, at less than 14 weeks' gestation, who spoke English or French. Exclusion criteria included foetal abnormalities, medical complications, illicit drug use during pregnancy, or other details previously described. Sample for this study: 601 mother–child dyads from follow-up phase (MIREC Child Development Plus); data from 5 mother–child dyads were excluded due to the mothers' declining prenatal and birth data collection (i.e. trimester fluoride exposures, demographic information, covariates, and offspring date of birth), leaving 596	44% of pairs resident in fluoridated cities; no information on dose or duration of exposure	56% of pairs resident in non-fluoridated cities; no information on duration	IQ	596 mother–child pairs with fluoride intake, IQ, and complete demographic and covariate data (non-fluoridated: ~334; fluoridated: ~262)

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		mother–child dyads for full analytic sample.				
Ibarluzea et al. (2021)	Prospective cohort study Spain	Mother-child pairs recruited during first trimester of pregnancy from one referral hospital in Spain; women aged 16 years or over, with singleton pregnancy achieved without assisted reproduction techniques, planned birth in the referral hospital, no communication problems in Spanish or Basque	Areas with artificially fluoridated water, fluoride concentration CWF areas 0.81 (±0.15 SD) mg/L	Areas without artificially fluoridated water, fluoride concentration <0.1 mg/L	Neuropsychological development of children at 1 year (Bayley Scales of Infant Development) and 4 years (McCarthy Scales of Children's Abilities)	393 women with complete information, 316 children included at age 1, 248 children included at age 4
Neurodevelopmental disorders						
Barberio *et al.* (2017a) (CHMS) 🟢	Cross-sectional survey Canada	Population-based sample of Canadian children aged 3–12 years living in private households in the 10 provinces (subset for whom information on sources of fluoride exposure was available)	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3 of surveys. Spot urine samples were available for a subsample of the respondents for Cycles 2 and 3, as specific gravity-adjusted urinary fluoride (micromoles per litre (µmol/L)) and creatinine-adjusted urinary fluoride (micromoles per millimole (µmol/mmol)); duration of exposure not specified	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3 of surveys. Spot urine samples were available for a subsample of the respondents for Cycles 2 and 3, as specific gravity-adjusted urinary fluoride (µmol/L) and creatinine-adjusted urinary fluoride (µmol/mmol)	Learning disability diagnosis (attention deficit disorder, no hyperactivity/ADHD/dyslexia/other)	Fluoride subsample (Cycle 2: 1,120; Cycle 3: 1,101)

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Riddell *et al.* (2019) (CHMS) ●	Cross-sectional survey Canada	National sample of Canadian youth aged 6–17 years from a population-based sample of Canadian residents aged 3–79 years living in private households (subset for whom information on sources of fluoride exposure was available (i.e. urine samples and tap water samples))	Fluoride in urine and tap water as a continuous variable. CWF as a dichotomous variable with mixed areas excluded. Duration of exposure not specified.	Fluoride in urine and tap water as a continuous variable. CWF as a dichotomous variable with mixed areas excluded.	ADHD diagnosis and scores	Fluoride in urine (1,877) and tap water (710) as continuous variables. CWF is a dichotomous variable (total: 1,722 (932 fluoridated; 790 non-fluoridated)) with mixed areas excluded.

3.4.2 Quality assessment

A summary of the quality assessment of the seven papers that examined neuropsychological outcomes, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 15 (see Appendix M for the full quality assessment). Six papers received a rating of low quality and one received a rating of high quality; this was the prospective cohort study that found no association between fluoride exposure and IQ in childhood or adulthood [59]. This study collected data on which participants received fluoridated water in their home; however, it is not known how much fluoridated water, tea, or other foods they consumed.

There are major limitations in the conduct of the MIREC study, on which three papers [54–56] based their analyses. The limitations include a non-random sample of maternity hospitals from the 10 selected cities; a clustering approach to selecting the 10 cities themselves and one of the maternity hospitals within each city, which was not accounted for in the sample size calculation or CIs; and a substantial loss to follow-up (only 808 dyads of the original cohort of 2,001 were invited to participate in the study of IQ; 601 participated and 512 provided full data). (These details were confirmed to the review team by Dr Tye Arbuckle, one of the principal investigators on the MIREC study (Arbuckle, personal communication, September 2021)). In addition, blinding status was not reported for two of the papers and was reported as a proxy status in the third. These limitations indicate that it is not clear whom the study population represents as a result of the selection process and high loss to follow-up, and whether the CIs around the main outcomes are valid, as it does not take account of the cluster design effect (Arbuckle, personal communication, September 2021), which would lead to wider CIs affecting statistical significance.

The study by Ibarluzea *et al.* (2021) [38] suffers from similar substantial limitations. The paper does not report on the participation rate and sample size considerations. The study also had substantial attrition from the first trimester to follow-up at 4 years of age and recruited participants from only one public referral hospital in one Spanish region, giving rise to concerns about generalisability and potential homogeneity within the sample. A 2018 analysis [86] from the same cohort study found that levels of environmental chemicals in the placenta, such as lead and arsenic, were significant predictors of reduced IQ; it is possible that measurement of the placenta, rather than maternal urine and cord blood, may have accounted more comprehensively for confounding chemical exposures and reduced the risk of residual confounding. This is of particular importance in this region, where mercury exposure is a known issue.

For Barberio *et al.*'s (2017a) [49] and Riddell *et al.*'s (2019) [50] cross-sectional survey analyses, the main limitation was a lack of reporting by these paper authors rather than limitations in the CHMS's sampling and analysis methodology. Cross-sectional surveys are useful for estimating prevalence, planning and evaluating health services, and identifying theoretical relationships; however, they cannot prove or disprove causality, as they collect exposure and outcome information at the same time.

Control for confounding factors, particularly in relation to neurotoxic environmental substances, is of particular importance. Lead was controlled for in four of the seven studies, while arsenic was controlled for in three.

Table 15 Summary of quality assessment ratings for studies examining neuropsychological outcomes

Item	Broadbent *et al.* (2015)	Green *et al.* (2019) (MIREC) 🔴	Till *et al.* (2020) (MIREC) 🔴	Farmus *et al.* (2021) (MIREC) 🔴	Ibarluzea *et al.* (2021)	Barberio *et al.* (2017a) (CHMS) 🟢	Riddell *et al.* (2019) (CHMS) 🟢
3. Was the participation rate of eligible persons at least 50%?	Yes	No	No	No	Not reported	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	No	No	No	No	No	No	No
13. Was loss to follow-up after baseline 20% or less?	Yes	No	No	No	No	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Almost all	Partial	Almost all	Almost all	Limited	Limited
Quality rating	3.0 (moderate)	2.0 (low)	1.5 (low)	2.0 (low)	2.0 (low)	2.0 (low)	2.0 (low)

3.4.3 Findings

A summary of the findings of the studies that examined neuropsychological outcomes, including IQ and neurodevelopmental disorders, is presented in Table 16 (see Appendix N for full findings). The findings are presented grouped by study, rather than by outcome, due to the interrelatedness of the findings within each study, particularly in relation to IQ. For example, as verbal IQ scores are correlated with FSIQ, these scores should be interpreted together and not independently of one another.

The prospective cohort study by Broadbent *et al.* (2015) [59] found no association between preschool fluoride exposure and FSIQ at the ages of 7–13 years or at the age of 38 years, nor did the study identify associations between fluoride exposure and IQ index scores for verbal comprehension, perceptual reasoning, working memory, and processing speed at the age of 38 years. IQ scores at the ages of 7–13 years were collected at the ages of 7, 9, 11, and 13 years; the IQ scores were averaged into one measure and then standardised. The analysis did not control for environmental neurotoxic substances (e.g. lead, arsenic).

Three papers analysed data from 601 of the 2,001 mother–child pairs enrolled in the MIREC longitudinal cohort study (Green *et al.* (2019) [54], Till *et al.* (2020) [55], and Farmus *et al.* (2021) [56]). These papers examined the influence of fluoride exposure on IQ at the ages of 3–4 years.

Green *et al.* (2019) [54] found evidence that, following adjustment, higher levels of maternal urinary fluoride were associated with lower FSIQ scores in boys but not in girls. Similar associations were found for fluoride intake and fluoride concentration in tap water among both boys and girls. It must be noted that these models included two boys with FSIQ scores lower than 60; it is highly likely that these scores are attributable to some neurodevelopmental condition, congenital abnormality, or birth injury and not to fluoride exposure. However, the association between higher maternal urinary fluoride and lower FSIQ remained statistically significant when the analysis was run excluding these two outliers. The model also remained statistically significant for this group in separate sensitivity analyses that additionally controlled for lead, mercury, perfluorooctanoic acid, arsenic, manganese, and second-hand smoke exposure.

The Green *et al.* (2019) [54] paper found no association between change in maternal fluoride, fluoride intake, or water fluoride concentration and verbal IQ. The paper found that higher maternal specific gravity-adjusted urinary fluoride was associated with lower performance IQ scores among boys. No difference was observed among girls. Higher fluoride concentration in tap water was also associated with lower performance IQ scores. No effect of change in fluoride intake was observed.

Till *et al.* (2020) [55] found that for children who were formula-fed as infants, higher water fluoride concentration was associated with lower FSIQ scores at the ages of 3–4 years; however, this association was no longer significant when maternal urinary fluoride was controlled for. The association was also rendered non-significant when the analysis was run without the two male outliers described above. No negative relationship between IQ and fluoridation status was observed for children who were breastfed as infants.

The same paper [55] found no association between changes in water fluoride concentration and verbal IQ at the ages of 3–4 years for either the breastfed or formula-fed groups. However, water fluoride concentration was associated with a lower performance IQ score at the ages of 3–4 years for both breastfed and formula-fed groups, which remained statistically significant after controlling for maternal urinary fluoride and after the removal of the two male outliers described above. The analysis did not control for the influence of environmental neurotoxic substances (e.g. lead, arsenic), although the data were available in the MIREC dataset.

Farmus *et al.* (2021) [56] found a differential influence of fluoride on IQ for different exposure periods: during pregnancy, during infancy, and during early childhood. Associations were most apparent for performance IQ; across both boys and girls, prenatal exposure to fluoride had the strongest negative influence on performance IQ and a somewhat weaker but still significant influence in infancy. The strongest associations in boys and girls were with prenatal exposure and exposure during infancy respectively, such that a 0.5 mg/L increase in maternal urinary fluoride was associated with a 4-point decrement in performance IQ for boys, and a 0.1 mg/L increase in infant fluoride intake was associated with a 2-point decrement in performance IQ for girls. The associations were weaker with FSIQ for each exposure period (during pregnancy, infancy, and early childhood), and were non-significant for verbal IQ. The analysis did not control for the influence of environmental neurotoxic substances; however, the authors state that preliminary analyses that did control for the influence of lead, mercury, perfluorooctanoic acid, and arsenic did not have an appreciable effect on their estimates.

The findings of Ibarluzea *et al.* (2021) [38] – specifically, the findings from the analysis controlling for mercury in umbilical cord blood, which we present here – are contradictory to the findings from the analyses of the MIREC study data. At the age of 4 years, higher maternal urinary fluoride during pregnancy was associated with higher scores on all cognitive scales for boys. However, once mercury in umbilical cord blood was accounted for in the analysis, only the associations for higher verbal and cognitive ability scores remained significant. This study found no association between maternal urinary fluoride during pregnancy and neuropsychological developmental scores at the age of 1 year. Additional analyses controlled for the influence of environmental neurotoxic substances, including arsenic, manganese, and lead, although only mercury had a substantial impact on the findings.

The study authors also carried out straightforward comparisons of scores between those living in fluoridated and non-fluoridated areas, finding a difference only for the numeric scale, which, again, favoured those children living in the fluoridated area. However, it should be noted that this dichotomous "fluoridated yes/no" variable renders the analysis ecological with respect to exposure, which is more vulnerable to confounding (e.g. lead and arsenic industrial pollution). This finding should also not be overinterpreted; the numeric scale is only one of four subdomains and the general cognitive index is a psychometrically stronger outcome.

Two cross-sectional papers presented evidence on the association between fluoride and neurodevelopmental disorders, both based on data from Statistics Canada's CHMS.

The first paper, by Barberio *et al.* (2017a) [49], found no association between reported diagnosis of ADHD, ADHD without hyperactivity (attention deficit disorder (ADD)), or learning disability, and any measure of fluoride exposure following adjustment for covariates.

However, later analysis of the same dataset in a separate paper (Riddell *et al.* (2019) [50]) found that a higher concentration of fluoride in tap water was associated with higher odds of ADHD diagnosis. The paper by Riddell *et al.* also found that for older youth (at the 75th percentile for age), living in a fluoridated area was associated with significantly higher odds of ADHD diagnosis and increased scores on a measure of hyperactivity/inattentiveness. The study also found an interaction between age and hyperactivity scores, such that the associations between: (a) higher hyperactivity/inattention scores and higher concentrations of fluoride in tap water, and (b) higher scores and living in a fluoridated area, were statistically significant among older youth but not younger youth. The clinical significance of this finding is not clear and interpretation is hampered by the relatively small number of participants with ADHD in the sample. However, it is possible that the pattern reflects the fact that ADHD is more likely to be diagnosed in older school-aged children than preschool-aged children. Thus, the different findings between the Barberio *et al.* analysis (age 3–12 years versus Riddell *et al.* (age 6–17 years) may reflect the greater

sensitivity in detecting an association for older youth, who have had both longer exposure to fluoride and a higher probability of having a diagnosis of ADHD than younger children. The Barberio *et al.* analysis controlled for no environmental neurotoxic substances, while the Riddell *et al*. analysis controlled for blood lead.

In summary, the three studies investigating the influence of fluoride on IQ have mixed findings; that is, two found generally no association, and one found a negative association. The positive associations between fluoride exposure and some measures of IQ in the study by Ibarluzea *et al.* should not be interpreted as evidence for a beneficial effect of CWF on IQ, and the study authors do not make such a claim. Two of the three cohort studies had high loss to follow-up (the MIREC study and the study presented by Ibarluzea *et al.* [38]) and the MIREC longitudinal cohort study has methodological issues that call into question the validity of the findings presented in these papers [54–56]. Studies analysing data from the CHMS present conflicting findings with respect to diagnosis of ADHD, and one analysis demonstrated stronger associations between fluoride exposure and hyperactivity for older youth, a pattern that may be explained by differences in diagnostic sensitivity across childhood and adolescence. A high-quality prospective longitudinal study based on individual-level exposures, taking account of all potential confounding factors, effect modifiers, and cluster design effect, is required in order to strengthen the evidence base on neuropsychological outcomes.

Table 16 Summary of findings for studies examining neuropsychological outcomes

Outcome	Comparisons	Method of measurement	Summary of findings
Broadbent *et al.* (2015)			
FSIQ at the ages of 7-13 years (measured at 4 ages and averaged)	Resident vs. never lived in CWF area	FSIQ, Wechsler Intelligence Scale for Children Revised (standardised to mean: 100; SD: 15)	No association between preschool fluoride exposure and IQ in childhood
FSIQ at the age of 38 years	Resident vs. never lived in CWF area	FSIQ, Wechsler Adult Intelligence Scale, Fourth Edition (standardised to mean: 100; SD ±15)	No association between preschool fluoride exposure and IQ in adulthood
Verbal comprehension index at the age of 38 years	Resident vs. never lived in CWF area	Verbal Comprehension Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)	No association between preschool fluoride exposure and scores in childhood
Perceptual reasoning index at the age of 38 years	Resident vs. never lived in CWF area	Perceptual Reasoning Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)	No association between preschool fluoride exposure and scores in childhood
Working memory index at the age of 38 years	Resident vs. never lived in CWF area	Working Memory Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)	No association between preschool fluoride exposure and scores in childhood
Processing speed index at the age of 38 years	Resident vs. never lived in CWF area	Processing Speed Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)	No association between preschool fluoride exposure and scores in childhood
Green *et al.* (2019) (MIREC) 🔴			
FSIQ at the ages of 3–4 years	Maternal urinary fluoride as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	1 mg increase in maternal urinary fluoride associated with a 4.49-point (−8.38 to −0.60) lower FSIQ score in boys; no association for total sample or for girls
FSIQ at the ages of 3–4 years	Maternal fluoride intake as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	1 mg increase in fluoride intake associated with a 3.66-point (−7.16 to −0.15) lower FSIQ score among boys and girls
FSIQ at the ages of 3–4 years	Water fluoride concentration as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	1 mg increase in water fluoride concentration associated with a 5.29-point (−10.39 to −0.19) lower FSIQ score among boys and girls
Verbal IQ at the ages of 3–4 years	Maternal urinary fluoride as predictor for total sample, boys, and girls	Verbal IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	No association between change in maternal urinary fluoride and verbal IQ
Verbal IQ at the ages of 3–4 years	Maternal fluoride intake as	Verbal IQ, Wechsler Preschool and Primary Scale	No association between change in fluoride intake and verbal IQ

Outcome	Comparisons	Method of measurement	Summary of findings
	predictor for total sample, boys, and girls	of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	
Verbal IQ at the ages of 3–4 years	Water fluoride concentration as predictor for total sample, boys, and girls	Verbal IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	No association between change in water fluoride concentration and verbal IQ
Performance IQ at the ages of 3–4 years	Maternal urinary fluoride as predictor for total sample, boys, and girls	Performance IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	Increase of 1 mg/L maternal urinary fluoride associated with a 4.63-point lower performance IQ score in boys, but no difference in girls
Performance IQ at the ages of 3–4 years	Maternal fluoride intake as predictor for total sample, boys, and girls	Performance IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	No effect of change in fluoride intake on performance score
Performance IQ at the ages of 3–4 years	Water fluoride concentration as predictor for total sample	Performance IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).	Increase of 1 mg/L water fluoridation associated with a 13.79-point (18.82–7.28) lower performance score for total sample
Till *et al.* (2020) (MIREC) 🔴			
FSIQ at the ages of 3–4 years	Water fluoride concentration as predictor for children who were formula-fed or breastfed as infants	FSIQ score differences, Wechsler Preschool and Primary Scale of Intelligence, Third Edition using USA population-based normative data (mean: 100; SD: 15).	Increase of 0.5mg /L in fluoride concentration associated with lower FSIQ scores for formula-fed group only (4.4 points lower (-8.34–0.46), no difference observed for breastfed group. This difference was no longer significant when maternal urinary fluoride was controlled for or after removal of IQ outliers.
Verbal IQ at the ages of 3–4 years	Water fluoride concentration as predictor for children who were formula-fed or breastfed as infants	Verbal Scale IQ score differences, Wechsler Preschool and Primary Scale of Intelligence, Third Edition using USA population-based normative data (mean: 100; SD: 15).	Water fluoride concentration not associated with changes in verbal IQ for either formula-fed or breastfed groups, remaining statistically non-significant when controlling for maternal urinary fluoride and following removal of IQ outliers
Performance IQ at the ages of 3–4 years	Water fluoride concentration as predictor for children who were formula-fed or breastfed as infants	Performance IQ score differences, Wechsler Preschool and Primary Scale of Intelligence, Third Edition using USA population-based normative data (mean=100, SD=15)	Increase of 0.5mg /L in water fluoride concentration significantly associated with lower performance IQ in both formula-fed (9.26 points lower (13.77–4.76)) and breastfed (6.19 points lower (10.45–

Outcome	Comparisons	Method of measurement	Summary of findings
			1.94)) groups, remaining statistically significant when controlling for maternal urinary fluoride
Farmus *et al.* (2021) (MIREC) 🔴			
FSIQ at the ages of 3–4 years	Standardised maternal urinary fluoride as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	Association between prenatal exposure to fluoride and FSIQ for boys only; no longer significant following removal of influential dyads in sensitivity analysis
FSIQ at the ages of 3–4 years	Standardised infant fluoride intake as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	No association between exposure to fluoride during infancy and FSIQ
FSIQ at the ages of 3–4 years	Standardised child urinary fluoride as predictor for total sample, boys, and girls	FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	No association between childhood exposure to fluoride and FSIQ
Verbal IQ at the ages of 3–4 years	Standardised maternal urinary fluoride as predictor for total sample, boys, and girls	Verbal Scale IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	No association between prenatal exposure to fluoride and verbal IQ
Verbal IQ at the ages of 3–4 years	Standardised infant fluoride intake as predictor for total sample, boys, and girls	Verbal Scale IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms;	No association between exposure to fluoride during infancy and verbal IQ
Verbal IQ at the ages of 3–4 years	Standardised child urinary fluoride as predictor for total sample, boys, and girls	Verbal Scale IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	No association between childhood exposure to fluoride and verbal IQ
Performance IQ at the ages of 3–4 years	Standardised maternal urinary fluoride as predictor for total sample, boys, and girls	Performance Scale IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	Association between prenatal exposure to fluoride and performance IQ for boys and for total sample, such that an increase of 0.5 mg/L in maternal urinary fluoride was associated with a 4.02-point (6.15–1.89) lower performance IQ score for boys and a 3.15-point (4.85–1.44) lower performance IQ score for total sample
Performance IQ at the ages of 3–4 years	Standardised infant fluoride	Performance Scale IQ, Wechsler Preschool and	Association between exposure to fluoride during infancy and

Outcome	Comparisons	Method of measurement	Summary of findings
	intake as predictor for total sample, boys, and girls	Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	performance IQ for girls and for total sample, such that an increase of 0.5 mg/L in maternal urinary fluoride was associated with a 2.03-point (−3.43 to −0.63) lower performance IQ score for girls and a 1.58-point (2.59–0.57) lower performance IQ score for total sample
Performance IQ at the ages of 3–4 years	Standardised child urinary fluoride as predictor for total sample, boys, and girls	Performance Scale IQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III; Canadian norms)	No association between childhood exposure to fluoride and performance IQ
Ibarluzea *et al.* (2021)			
Neuropsychological development at the age of 1 year	Maternal urinary fluoride as predictor for total sample, boys, and girls	Standardised scores on Bayley Mental Development Index (81tandardized to mean: 100; SD: 15)	No significant association between maternal fluoride during pregnancy and scores on Bayley Mental Development Index at the age of 1 year
Neuropsychological development at the age of 4 years: General cognitive	Maternal urinary fluoride as predictor for total sample, boys, and girls	Standardised scores on McCarthy Scales of Children's Abilities (General Cognitive scale), adapted to Spanish population (standardised to mean: 100; SD: 15)	1 mg increase in maternal urinary fluoride across whole pregnancy associated with a 10.54-point (0.19 to 20.89) higher general cognitive score in boys when adjusted for blood cord mercury levels; no association for girls. 1mg increase in maternal urinary fluoride in third trimester associated with a 8.15-point (0.69 to 15.61) higher general cognitive score in boys when adjusted for blood cord mercury levels; no association for girls.
Neuropsychological development at the age of 4 years: Verbal	Maternal urinary fluoride as predictor for total sample, boys, and girls	Standardised scores on McCarthy Scales of Children's Abilities (Verbal scale), adapted to Spanish population (81tandardized to mean: 100; SD: 15)	1mg increase in maternal urinary fluoride across whole pregnancy associated with a 9.74-point (1.75 to 17.74) higher verbal score in boys when adjusted for blood cord mercury levels; no association for girls.
Neuropsychological development at the age of 4 years: Performance	Maternal urinary fluoride as predictor for total sample, boys, and girls	Standardised scores on McCarthy Scales of Children's Abilities (Performance scale), adapted to Spanish population (standardised to mean: 100; SD: 15)	No significant association between maternal fluoride during pregnancy and performance scores at the age of 4 years when adjusted for blood cord mercury levels
Neuropsychological development at the	Maternal urinary fluoride as	Standardised scores on McCarthy Scales of	No significant association between maternal fluoride

Outcome	Comparisons	Method of measurement	Summary of findings
age of 4 years: Numeric	predictor for total sample, boys, and girls	Children's Abilities (Numeric scale), adapted to Spanish population (standardised to mean: 100; SD: 15)	during pregnancy and numeric scores at the age of 4 years when adjusted for blood cord mercury levels
Neuropsychological development at the age of 4 years: Memory	Maternal urinary fluoride as predictor for total sample, boys, and girls	Standardised scores on McCarthy Scales of Children's Abilities (Memory scale), adapted to Spanish population (standardised to mean: 100; SD: 15)	No significant association between maternal fluoride during pregnancy and memory scores at the age of 4 years when adjusted for blood cord mercury levels
Barberio *et al.* (2017a) (CHMS) 🟢			
Learning disability at the ages of 3–12 years	Outcome regressed on urinary fluoride and fluoride concentration in tap water	Parental- or self-reported diagnosis of a learning disability	Reported learning disability diagnosis not significantly associated with creatinine-adjusted urinary fluoride, specific gravity-adjusted urinary fluoride, or fluoride concentration of tap water in adjusted or unadjusted models
ADHD diagnosis in childhood/ adolescence	Outcome regressed on urinary fluoride	Parental- or self-reported diagnosis of ADHD at the ages of 3–12 years	Reported diagnosis of ADHD not significantly associated with creatinine-adjusted urinary fluoride, specific gravity-adjusted urinary fluoride
ADD diagnosis at the ages of 3–12 years	Outcome regressed on urinary fluoride	Parental- or self-reported diagnosis of ADD (no hyperactivity) at the ages of 3–12 years	Reported diagnosis of ADD significantly associated with creatinine-adjusted urinary fluoride, such that those with higher creatinine-adjusted urinary fluoride had lower odds of reporting ADD; however, association was reduced to non-significance in the adjusted model
Riddell *et al.* (2019) (CHMS) 🟢			
ADHD diagnosis in childhood/ adolescence	Outcome regressed on urinary fluoride	Physician-made diagnosis of ADHD at the ages of 6–17 years	Urinary fluoride did not significantly predict ADHD diagnosis
ADHD diagnosis in childhood/ adolescence	Outcome regressed on fluoride concentration in tap water	Physician-made diagnosis of ADHD at the ages of 6–17 years	1 mg/L increase in tap water fluoride was associated with 6.1 (1.60–22.8) times higher odds of ADHD diagnosis; no interaction with age or sex
ADHD diagnosis in childhood/ adolescence	Outcome regressed on CWF status (binary)	Physician-made diagnosis of ADHD at the ages of 6–17 years	Significant interaction between age and CWF status, such that for older youth (at the 75th percentile for age), predicted odds of ADHD diagnosis were 2.84 (1.40–5.76) times higher among youth in a fluoridated region than in a non-fluoridated region; no difference across regions in odds for youth at the 25th percentile for age

Outcome	Comparisons	Method of measurement	Summary of findings
Hyperactivity/ inattention scores	Outcome regressed on urinary fluoride	Scores, hyperactivity/inattention scale on the Strengths and Difficulties Questionnaire	Urinary fluoride did not significantly predict hyperactivity/inattention scale scores
Hyperactivity/ inattention scores	Outcome regressed on fluoride concentration in tap water	Scores, hyperactivity/inattention scale on the Strengths and Difficulties Questionnaire	Significant interaction between age and hyperactivity/inattention scale scores;1 mg/L increase in tap water fluoride was associated with a 1.52 (0.23–2.80) increase in scores for youth at 75th percentile for age; not significant for youth at 25th percentile
Hyperactivity/ inattention scores	Outcome regressed on CWF status (binary)	Scores, hyperactivity/inattention scale on the Strengths and Difficulties Questionnaire	Significant interaction between age and hyperactivity/inattention scale scores; for youth at 75th percentile for age, living in a fluoridated region was associated with a 0.7-point higher score (0.34–1.06); no association between CWF status and scores for youth at 25th percentile

3.5 Cancer

A possible link between water fluoridation and higher cancer mortality was claimed in the 1970s, and the possibility raised health concerns and heightened controversy surrounding the practice of CWF [25]. In 1977, Yiamouyiannis and Burk [87] reported that cancer mortality was higher in areas with artificially fluoridated drinking water than in non-fluoridated areas. These findings were subsequently refuted by other investigators who identified problems with the study's research methodology [88].

An association between fluoride and bone cancers is theoretically plausible. It is known that on a molecular level, fluoride is deposited on bone tissue, and this has a mitogenic effect on osteoblasts – in other words, it triggers cell division in bone cells [47]. An association with bladder cancer is also theoretically plausible; fluoride is excreted through the bladder, so the bladder lining is exposed to relatively high concentrations of fluoride.

Twelve papers [20,41,47,48,58,61,62,65,68,70–72] presented data from 11 studies examining the association between fluoridated water and cancer incidence. Bone cancers (including osteosarcoma and Ewing sarcoma) were the subject of interest in 11 papers (10 studies), while secondary bone cancer, bladder cancer, and all cancers were each examined by 1 paper.

3.5.1 Bone cancers

Eleven papers [20,41,47,48,58,61,65,68,70–72] presented data from 10 studies examining the association between bone cancers and fluoridated water status. Osteosarcoma was the cancer examined in 10 papers (9 studies), bone cancers in general were examined by 2 papers, and Ewing sarcoma was examined by 1 paper.

3.5.1.1 Study characteristics

A summary of the characteristics of the 11 papers that examined bone cancers is presented in Table 17 (see Appendix L for full study characteristics). Eight were ecological papers and three were case-control papers.

Six of the eight ecological studies examined differences in incidences of cancer between fluoridated and non-fluoridated areas [20,41,61,65,70,72], while one examined differences according to the extent of implementation of fluoridation [68] and one examined differences in cancer incidence according to concentration of fluoride [58]. In the case of Cohn (1992) [61], we have extracted only the findings pertaining to the main seven-county area described in the paper, not the subset analysis of a three-county area or the separate analysis of a naturally fluoridated county.

The three case-control papers (based on two studies) examined exposure to fluoridated water based on residential history using osteosarcoma patients and their matched controls [47,48,71]. It is important to note that the 2020 case-control paper by Kim *et al.* [48] is an update of the 2006 case-control study by Bassin *et al.* [47], and incorporates a broader population and other variables of interest (e.g. bottled water); these papers are based on the Harvard Fluoride Osteosarcoma Study.

Case-control studies are useful for investigating rare diseases such as osteosarcoma. Of note is the fact that case-control studies and retrospective cohort studies suffer from recall bias, which is a problem for all the included case-control studies, but it is more severe in the case of Bassin *et al.*, due to the long delay between diagnosis and the data collection on exposure.

Two papers [47,71] excluded patients with a history of kidney dialysis, as such patients chose to drink deionised water for medical reasons.

Table 17 Summary of study characteristics for studies examining bone cancers

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Osteosarcoma						
Hrudey *et al.* (1990)	Ecological or correlational study Canada	Cases of osteosarcoma recorded by the Alberta Cancer Board registry, Canada	Incidence of osteosarcoma in Edmonton, Alberta, Canada, where water has been fluoridated to 1.0 mg/L since 1967, recorded by the Alberta Cancer Board	Incidence of osteosarcoma recorded by the Alberta Cancer Board in Calgary, where water is not fluoridated (natural levels average 0.3 mg/L)	Osteosarcoma	Edmonton (fluoridated) 26 cases, Calgary (non-fluoridated) 29 cases
Mahoney *et al.* (1991)	Ecological or correlational study USA	Bone cancers recorded by the New York State Cancer Registry, USA, primary malignancies of the bone	Primary analysis: fluoridated areas in New York State, exclusive of New York City (due to lack of outcome data before 1973). Additional analysis: fluoridated counties located within standard metropolitan statistical areas (i.e. urbanised areas) and fluoridated counties not located within standard metropolitan statistical areas. Level of fluoridation not specified, and duration of exposure not specified.	Primary analysis: non-fluoridated areas in New York State, including some metropolitan areas that have maintained non-fluoridated water supplies. Additional analysis: three urbanised metropolitan areas that have maintained non-fluoridated water supplies.	Bone cancer, including osteosarcoma	Bone cancer cases (n=228) and osteosarcoma cases (n= 108)
McGuire *et al.* (1991)	Matched case-control study USA	Cases: patients diagnosed with osteosarcoma between 1980 and 1990, aged under 40 years at diagnosis, identified from the University of Iowa Cancer Registry and the medical records of the Division of Orthopaedics, St 'Joseph's Hospital in Omaha, Nebraska, USA, excluding patients with prediagnosis history of radiation therapy or	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	22 cases

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		kidney disease. Hospital-based controls from orthopaedics department matched by age, gender, and county of residence at time of diagnosis.				
Cohn (1992)	Ecological or correlational study USA	Cases of osteosarcoma in those aged under 20 years in seven central New Jersey, USA counties, compiled from the New Jersey Cancer Registry between 1979 and 1987; population data gathered from 1980 United States Census on a municipality level	Fluoridated municipalities where >85% of the population was supplied with fluoridated water from at least the early 1970s to at least 1987; United States standard target fluoride level was 0.7–1.2 ppm at time of study	Non-fluoridated municipalities where <10% of the population was supplied with fluoridated water; levels not specified	Osteosarcoma	Under 20 years of age: Population: 721,347 Cases: 30
Bassin *et al.* (2006) (Harvard) 🔵	Case-control study USA	Cases of osteosarcoma diagnosed before the age of 20 years in 11 hospitals across the USA (excluding those aged over 40 years, or with a history of radiation therapy or renal dialysis); controls were patients of the same hospitals' orthopaedics departments seen within ±6 months of cases' diagnosis and matched for age (±5 years), gender, distance from hospital, same exclusion criteria	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	Cases 103 and Controls 215

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Comber *et al.* (2011)	Ecological or correlational study Ireland	Cases of osteosarcoma recorded in the Northern Ireland Cancer Registry and the National Cancer Registry of Ireland between 1994 and 2006. Population denominator assigned by electoral district – mid-year population estimates for Northern Ireland, provided annually by the Northern Ireland Statistics and Research Agency; Republic of Ireland census data for 1996, 2002, and 2006 interpolated for intervening years.	Fluoridated regions (non-rural Republic of Ireland electoral divisions), levels not specified, targets 0.6 – 1.0 ppm during this time period; duration not specified; water fluoridation signed into law in 1960	Non-fluoridated regions (rural Republic of Ireland electoral divisions and Northern Ireland); levels not specified, generally not more than 0.3 ppm in Ireland	Osteosarcoma	183 cases
National Fluoridation Information Service (2013)	Ecological or correlational study New Zealand	Cases of osteosarcoma recorded in the New Zealand Cancer Registry between 2000 and 2008; rate per 1,000,000 population calculated but no information given on source of denominator data	Census area units served by CWF; levels not specified but generally fluoridated to 0.7–1.0 ppm in New Zealand; duration of exposure not specified	Census area units not served by CWF; levels not specified, generally not more than 0.3 ppm in New Zealand	Osteosarcoma	127 cases (fluoridated 58 and non-fluoridated 69)
Blakey *et al.* (2014)	Ecological or correlational study United Kingdom	Cases of osteosarcoma or Ewing sarcoma diagnosed in Great Britain from 1980 to 2005 drawn from population-based cancer registries. Denominator data derived from	Level of fluoride in water for census small area units in 2001; duration of exposure not specified	Level of fluoride in water for census small area units in 2001 (introduction data: optimal fluoride: 0.7–1.2 ppm; non-fluoridated: 0.3 ppm (confers no dental benefit))	Osteosarcoma and Ewing sarcoma	Osteosarcoma cases 2,566 and Ewing sarcoma cases 1,650

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		national decennial census data, population counts from previous censuses adjusted to be compatible with 2001 Census boundaries, census wards in England and Wales, and postcode sectors in Scotland.				
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Osteosarcoma	Population under 25 years: fluoridated areas 31,313,151, non-fluoridated areas 216,921,400 Cases under 25 years: fluoridated areas 148, non-fluoridated areas 949 Population 50 years and over: fluoridated areas 33,080,465, non-fluoridated areas 232,282,090 Cases 50 years and over: fluoridated areas 73, non-

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
						fluoridated areas 587.
Kim *et al.* (2020) (Harvard)	Case-control study USA	Cases: Phase 1: all patients diagnosed with osteosarcoma by participating departments in the USA from 1989 to 1993, histologically confirmed; Phase 2: cases of osteosarcoma identified by physicians in participating departments, diagnosed and treated with primary osteosarcoma confirmed by surgical pathology reports from 1994 to 2000. Controls: Phase 1: patients of record from 1989 to 1993 with other bone tumours or nonneoplastic conditions identified from same orthopaedic surgery departments as cases; Phase 2: patients from same departments as cases from 1994 to 2000 with newly diagnosed malignant bone tumours other than osteosarcoma (tumour controls) and benign tumours and nonneoplastic conditions (orthopaedic controls),	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	All 645; Cases 236; Controls 409;

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		including ambulatory orthopaedic patients.				
Ewing sarcoma						
Blakey *et al.* (2014)	Ecological or correlational study United Kingdom	Cases of osteosarcoma or Ewing sarcoma diagnosed in Great Britain from 1980 to 2005 drawn from population-based cancer registries. Denominator data derived from national decennial census data, population counts from previous censuses adjusted to be compatible with 2001 Census boundaries, census wards in England and Wales, and postcode sectors in Scotland.	Level of fluoride in water for census small area units in 2001; duration of exposure not specified	Level of fluoride in water for census small area units in 2001 (introduction data: optimal fluoride: 0.7–1.2 ppm; non-fluoridated: 0.3 ppm (confers no dental benefit))	Osteosarcoma and Ewing sarcoma	Osteosarcoma cases 2,566 and Ewing sarcoma cases 1,650
Bone cancer						
Mahoney *et al.* (1991)	Ecological or correlational study USA	Bone cancers recorded by the New York State Cancer Registry, USA, primary malignancies of the bone	Primary analysis: fluoridated areas in New York State, excluding New York City (due to lack of outcome data before 1973). Additional analysis: fluoridated counties located within standard metropolitan statistical areas (i.e. urbanised areas) and fluoridated counties not located within standard metropolitan statistical areas. Level of fluoridation not specified; duration of exposure not specified.	Primary analysis: non-fluoridated areas in New York State, including some metropolitan areas that have maintained non-fluoridated water supplies. Additional analysis: three urbanised metropolitan areas that have maintained non-fluoridated water supplies.	Bone cancer, including osteosarcoma	Bone cancer cases (n=228) and osteosarcoma cases (n= 108)

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in 11 areas, withdrawn in 2004 in all areas	7 areas did not receive CWF.	Bone cancer	Populations: CWF 4,406,021 and Non-CWF 2,270,959

3.5.1.2 Quality assessment

A summary of the quality assessment of the eight ecological papers that examined bone cancer, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 18 (see Appendix M for the full quality assessment). All eight received a rating of moderate quality. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 18 Summary of quality assessment ratings for ecological studies examining bone cancers

Item	Hrudey *et al.* (1990)	Mahoney *et al.* (1991)	Cohn (1992)	Comber *et al.* (2011)	National Fluoridation Information Service (2013)	Blakey *et al.* (2014)	Young *et al.* (2015)	Lee *et al.* (2020)
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Limited	Limited	Limited	Limited	Limited	Limited
Quality rating	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)

A summary of the quality assessment of the three case-control papers that examined bone cancer, using the NHLBI quality assessment tool for case-control studies, is reported in Table 19 and the full quality assessment is provided in Appendix M. One paper was rated high quality and two were rated low quality. The three case-control papers are based on individual experiences of exposures and outcomes, and can contribute to evidence for causality.

Case-control studies are useful to investigate rare diseases, such as osteosarcoma. It is important to note that case-control studies are retrospective in nature; the outcome (i.e. presence or absence of the condition of interest) is already known at the outset and is, indeed, the starting point for the selection of participants. In this way, the investigation proceeds temporally in the reverse order to other epidemiological study designs, from effect (disease) to cause (antecedent exposure); both the exposure and the disease have already occurred when participants are selected for inclusion. Although highly efficient and well suited for studying uncommon diseases like osteosarcoma, case-control studies are therefore at particular risk for various forms of bias in both the selection of participants and the reporting or recall of information about their exposure [39]. Recall bias is a particular concern in the case of Bassin *et al.* [47] due to the long delays between diagnosis and the data collection on exposure. In addition, blinding status was not reported.

Table 19 Summary of quality assessment ratings for case-control studies examining osteosarcoma

Item	McGuire *et al.* (1991)	Bassin *et al.* (2006) (Harvard)	Kim *et al.* (2020) (Harvard)
4. Did the authors include a sample size justification?	No	No	No
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes	Yes	Partly
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	Yes	Partly	Partly
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Partial	Partial	Partial
Quality rating	3.5 (high)	2.5 (low)	2.5 (low)

3.5.1.3 Findings

The summary findings of the eight ecological studies that examined bone cancers (including osteosarcoma, Ewing sarcoma, and bone cancer generally) are presented in Table 20 (see Appendix N for full findings).

Seven ecological studies [20,41,58,61,65,68,70,72] examined osteosarcoma and water fluoridation. Six of these studies [20,41,58,65,68,70,72] found no difference in the incidence of osteosarcoma across fluoridated and non-fluoridated areas. One study [61] reported an elevated rate of osteosarcoma in 10–19-year-old males living in fluoridated areas; no other differences for any age group or sex were reported by this study.

One ecological study examined Ewing sarcoma, finding no difference in incidence rates of the disease according to fluoridation status [58].

Two ecological studies examined any bone cancers; neither found a difference in incidence rates of disease according to fluoridation status [68,70].

In summary, the evidence from these ecological studies does not suggest any association between CWF and the incidence of bone cancers.

Table 20 Summary of findings for ecological studies examining bone cancer (osteosarcoma, Ewing sarcoma, and any bone cancer)

Paper	Comparisons	Method of measurement	Summary of findings
Osteosarcoma			
Hrudey *et al.* (1990)	Fluoridated vs. non-fluoridated regions	Incidence of osteosarcoma recorded by the Alberta Cancer Board from 1970 to 1988	Similar incidence of osteosarcoma in fluoridated and non-fluoridated communities (no statistical comparison performed)
Mahoney *et al.* (1991)	Fluoridated vs. non-fluoridated regions for men and women aged under and over 30 years at time of diagnosis	Incidence of osteosarcoma recorded by the New York State Cancer Registry from 1975 to 1987	No difference in incidence between areas with and without fluoridated water for any age or sex groups
Cohn (1992)	Fluoridated vs. non-fluoridated regions for men and women across five age groups	Incidence of osteosarcoma in people aged under 20 years compiled from the New Jersey Cancer Registry between 1979 and 1987	Rate ratios elevated in fluoridated areas for males aged 10–19 years (relative rate ratio: 3.4 (1.8–6.0)). No difference in relative rate ratios for women or men in other age groups.
Comber *et al.* (2011)	Republic of Ireland fluoridated areas vs. all-Ireland non-fluoridated areas for men and women aged 0–24 years, and of all ages	Osteosarcoma incidence in Northern Ireland and Republic of Ireland, 1994–2006	No evidence of a significant association between water fluoridation and osteosarcoma incidence for any age or sex groups
National Fluoridation Information Service (2013)	Fluoridated vs. non-fluoridated regions for men and women across five age groups	Osteosarcoma incidence, diagnosed from 2000 to 2008	Osteosarcoma is extremely rare in New Zealand, with an average of 14.1 cases per year. No difference in rates of osteosarcoma cases between areas with and without CWF for both sexes (no statistical comparisons performed; descriptive statistics only).
Blakey *et al.* (2014)	Artificial fluoridation as binary variable	Osteosarcoma incidence, diagnosed from 1980 to 2005 in those aged 0–49 years	No association between artificial fluoridation (as a binary variable with adjustment for deprivation) and osteosarcoma
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	All cases in England diagnosed between 1995 and 2010 recorded in cancer registries	No difference in rate of osteosarcoma in all those aged under 25 years, males or females aged under 25 years, or in people aged over 50 years between fluoridated and non-fluoridated areas following adjustment for age, gender, deprivation, and ethnicity
Ewing sarcoma			

Paper	Comparisons	Method of measurement	Summary of findings
Blakey *et al.* (2014)	Artificial fluoridation as binary variable	Ewing sarcoma incidence, diagnosed from 1980 to 2005, in those aged 0–49 years	No association between artificial fluoridation (as a binary variable with adjustment for deprivation) and Ewing sarcoma
Bone cancer			
Mahoney *et al.* (1991)	Fluoridated vs. non-fluoridated regions for men and women aged under and over 30 years at time of diagnosis	Incidence of bone cancer recorded by the New York State Cancer Registry from 1975 to 1987	No difference in incidence between areas with and without fluoridated water for any age or sex groups
Lee *et al.* (2020)	Risk over time in fluoridated and non-fluoridated regions for total sample, men, and women	Incidence of bone cancer; data gathered from the National Health Insurance Service	Relative risks increased over time but did not increase in fluoridated regions compared with non-fluoridated regions

The findings of the three papers [47,48,71] based on two case-control studies that examined osteosarcoma are presented in Table 21.

The results from these three papers were varied. One study of 22 matched case-control pairs [71] found no negative association between exposure to fluoridated water 100% of the time and a diagnosis of osteosarcoma.

The 2006 paper by Bassin *et al.* [47], which included 103 cases and 215 controls, found that fluoride levels in drinking water (designated as 30–99% and 100% or over the local target level) before the age of 7 years was associated with a higher risk of osteosarcoma for males diagnosed with osteosarcoma before the age of 20 years. However, the 2020 paper by Kim *et al.* [48], which is an update of the 2006 Bassin *et al.* study and includes its participants, found that 236 osteosarcoma cases had lower odds than the 409 controls of having been exposed to fluoridated water for their whole lives. This updated paper found evidence that ever having lived in a fluoridated area had a significant protective effect against osteosarcoma for those who did not drink bottled water; this protective effect was not demonstrated for those who drank bottled water. The study also found evidence of a protective effect for those who had lived less than 50% of their lives in a fluoridated community compared with those who never had; however, there was no association for those who had lived more than 50% of their lives in a fluoridated community compared with those who never had.

Overall, any relationship between a diagnosis of osteosarcoma and exposure to CWF is unlikely, based on the evidence from case-control studies. Therefore, no relationship can be firmly established. A high-quality empirical study, taking account of all potential confounding exposures and effect modifiers, is required in order to strengthen the evidence base in this area.

Table 21 Summary of findings for case-control studies examining osteosarcoma

Paper	Comparisons	Method of measurement	Summary of findings
Osteosarcoma			
McGuire *et al.* (1991)	Pairs in which either the case or control had lived more than one-third of their lives with exposure to fluoride levels >0.7 ppm, had high average exposure to fluoride, or had lived more than one-third of the first 15 years of their lives with exposure to fluoride levels >0.7 ppm	Cases of osteosarcoma (before the age of 40 years), diagnosed from 1980 to 1990	No associations were significant; no evidence that exposure to fluoride is a risk factor for osteosarcoma
Bassin *et al.* (2006) (Harvard) 🔵	Less than 30% of target fluoride exposure at the age of 7 years vs. 30–99% of target fluoride exposure vs. at least 100% of target fluoride exposure for men and women	Cases of osteosarcoma (before the age of 20 years)	Fluoride level in drinking water before the age of 7 years associated with higher risk of osteosarcoma for men only (less than 30% target exposure; odds ratio: 3.36 (0.99–11.42) for 30–99% target exposure, 5.46 (1.50–19.90) for at least 100% target exposure); no association for women
Kim *et al.* (2020) (Harvard) 🔵	Ever lived vs. never lived in fluoridated area for participants who did and did not drink bottled water; 0% vs. <50% vs. >50% vs. 100% of life lived in fluoridated area	Incidence of osteosarcoma in participating departments from 1989 to 1993 and 1994 to 2000	Ever having lived in a fluoridated community, including those who did not drink bottled water, showed significant protective effect against osteosarcoma (odds ratio: 0.51 (0.31–0.84), *p*=0.008); protective effect not demonstrated for those who drank bottled water. Significant protective effect for those who had lived <50% of their lives in fluoridated area compared with those who never had (odds ratio: 0.41 (0.22–0.76)).

3.5.2 Other cancers

3.5.2.1 Study characteristics

The summary characteristics of the two ecological studies that examined other cancers are presented in Table 22 (see Appendix L for full study characteristics).

One study examined differences in incidences of any cancer and bladder cancer between fluoridated and non-fluoridated areas [41], while the other examined differences in secondary bone cancer incidences according to the extent of fluoridation implementation [62].

Table 22 Summary of study characteristics for studies examining other cancers

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Secondary bone cancer						
Crnosija *et al.* (2019)	Ecological or correlational study USA	Population: inpatient cancer patients drawn from the Statewide Planning and Research Cooperative System database run by the New York State Department of Health, USA; excluded those without a complete ZIP code, a patient ID code, or a New York State residency status, or who were aged under 18 years. Cases: presence or absence of the diagnosis code for secondary bone cancer (metastasis).	Limited variation in concentration of fluoride in fluoridated water supplies (45 counties received 0.7 mg/L; 2 counties 0.8 mg/L; 1 county 0.5 mg/L; and 1 county 0.4 mg/L); fluoridation was therefore evaluated by percentage of the population in each county that received public fluoridated water, and divided into three categories: <25%, 25–75%, and >75%; duration not specified.	Limited variation in concentration of fluoride in fluoridated water supplies (45 counties received 0.7 mg/L; 2 counties 0.8 mg/L; 1 county 0.5 mg/L; and 1 county 0.4 mg/L); fluoridation was therefore evaluated by percentage of the population in each county that received public fluoridated water, and divided into three categories: <25%, 25–75%, and >75%; duration not specified.	Secondary bone cancer	Not reported
Bladder cancer						
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Bladder cancer	Population: fluoridated areas 67,978,298, non-fluoridated areas 487,149,150 Cases: fluoridated areas: 11,327; non-fluoridated areas: 84,780
All cancers						

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	All cancers	Population: fluoridated areas 25,314,612, non-fluoridated areas 183,256,350 Cases: fluoridated areas: 131,288 non-fluoridated areas: 921,583

3.5.2.2 Quality assessment

A summary of the quality assessment of the ecological studies that examined other cancers, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 23 (see Appendix M for the full quality assessment), with separate assessments for the Young *et al.* analyses on bladder and all cancers to account for possible differences in as control for confounding, definition of exposure and outcome measures, and sample size. All three studies received A rating of moderate quality was given in all three cases. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 23 Summary of quality assessment ratings for ecological studies examining other cancers

Item	Crnosija *et al.* (2019)	Young *et al.* (2015) (bladder cancer)	Young *et al.* (2015) (all cancers)
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Unable to determine	Limited	Unable to determine
Quality rating	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)

3.5.2.3 Findings

The summary findings of the two ecological studies that examined secondary bone cancer, bladder cancer, and all cancers are presented in Table 24 (see Appendix N for full findings).

In one ecological study of cancer patients, no relationship was found between the percentage of county-level access to fluoridated water and the prevalence of secondary bone cancer [62].

The other ecological study, which was conducted in England, found a lower incidence of bladder cancer and all cancers in fluoridated areas (compared with non-fluoridated areas) after adjusting for confounders [41].

While these results suggest some possible protective effects of fluoridation against some forms of cancer, that association is based on a single ecological study and cannot be considered causal. A high-quality prospective longitudinal study based on individual-level exposures and taking account of all potential confounding factors is required in order to strengthen the evidence base on cancer.

Table 24 Summary of findings for ecological studies examining other cancers

Paper	Comparisons	Method of measurement	Summary of findings
Secondary bone cancer			
Crnosija *et al.* (2019)	<25% vs. 25–75% vs. 100% population in county with water fluoridation	County-level percentage of secondary bone cancer over cancer diagnosis	No relationship between county-level percentage of access to fluoridated water and prevalence of secondary bone cancer diagnosis among cancer patients
Bladder cancer			
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	All primary invasive bladder cancer cases in England diagnosed between 2000 and 2010 and recorded in cancer registries	Lower rates of bladder cancer in fluoridated areas (adjusted incidence rate ratio: 8.0 (−9.9 to −6.0), p<0.001)
All cancers			
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	All cancer cases in England (excluding non-melanoma skin cancer) diagnosed between 2007 and 2010 and recorded in cancer registries	All cancer incidence lower in fluoridated areas following adjustment for age, gender, and deprivation; however, this was not maintained when also adjusted for ethnicity.

3.6 Endocrine conditions

Concerns have been raised in relation to a possible relationship between endocrine gland dysfunction and fluoridated water. The glands that have been mentioned in this context are the pancreas (diabetes), the thyroid (goitre, hypothyroidism, and hyperthyroidism), the pituitary gland (hyperpituitarism and hypopituitarism), and the pineal gland (sleep disorders) [25]. Fluoride was used as a thyroid suppressant in the 1950s in order to treat hyperthyroidism, and is known to be associated with iodine deficiency [11]. The pineal gland is a small neuroendocrine organ near the centre of the brain but situated outside the blood–brain barrier, so it is exposed to fluoride in the bloodstream. Fluoride could theoretically accumulate in the pineal gland, leading to possible mineralisation of the gland, which may lead to changes in melatonin levels and possible sleep disruption [52].

Four papers [11,51–53], presenting data from two studies, examined the association between fluoridated water and endocrine conditions. Three papers [11,51,53] based on two studies examined a range of outcomes related to thyroid functioning, including incidence of diagnoses of thyroid disorders and thyroid-stimulating hormone (TSH) levels, while one paper [52] examined the incidence of sleep disturbances, which the study authors attribute to the functioning of the pineal gland.

3.6.1 Study characteristics

The summary characteristics of the four papers that examined endocrine conditions are presented in Table 25 (see Appendix L for full study characteristics). Three of these papers were based on a single cross-sectional study, and one paper was based on an ecological study.

One paper examined differences in the incidences of hypothyroidism between fluoridated and non-fluoridated areas [11], two papers examined the association between exposure to levels of fluoride and the outcomes of interest (TSH levels or sleep disturbance) [51,52], and one paper examined whether iodine deficiency modifies the impact of fluoride exposure on thyroid functioning [53].

It is important to note that three papers (Barberio *et al.* (2017b) [51], Cunningham *et al.* (2021) [52], and Malin *et al.* (2018) [53]) analyse data from the same dataset, Cycles 2 (2009–2011) and 3 (2012–2013) of

Statistics Canada's CHMS. Of these three papers, two (Barberio *et al.* (2017b) [51] and Malin *et al.* (2018) [53]) examined thyroid functioning – specifically TSH levels – as an outcome of interest, and the third paper (Cunningham *et al.* (2021) [52] examined aspects of sleep as an outcome of interest.

Table 25 Summary of study characteristics for studies examining endocrine conditions

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Thyroid functioning						
Peckham *et al.* (2015)	Ecological or correlational study England	Cases of hypothyroidism recorded by general practices in England in 2012 and 2013 using the Quality and Outcomes Framework	Levels of fluoride concentration assigned to general practices based on postcodes; target level 1 ppm (2012 data); maximum level >0.3 mg/L for all practices in the West Midlands; duration: <1 year	Levels of fluoride concentration assigned to general practices using postcodes; target level not specified (2012 data); maximum level ≤0.3 mg/L for all general practices in Greater Manchester	Hypothyroidism	946 general practices recruited; number included in CWF analysis not clear
Barberio *et al.* (2017b) (CHMS)	Cross-sectional survey Canada	Population-based sample of Canadian residents aged 3–79 years living in the 10 provinces; excluded those in the 3 territories (Cycle 3 only), those who live on reserves and other Aboriginal settlements in the provinces, institutionalised residents, full-time members of the Canadian Armed Forces, and residents of certain remote regions (subset for whom information on sources of fluoride exposure was available)	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3; spot urine samples were available for a subsample of the respondents for Cycles 2 and 3; duration of exposure not specified	Individual-level fluoride from urine and tap water samples	Thyroid functioning and/or diagnosis	Cycle 2 – fluoride urine subsample: 2,530; Cycle 3 – fluoride urine subsample: 2,671
Malin *et al.* (2018) (CHMS)	Cross-sectional survey Canada	Participants in the CHMS: randomly selected individuals aged 3–79 years living in 16 sites across the 10 Canadian provinces; excluded residents of the 3 territories, reserves, and Aboriginal settlements; full-time	Urinary fluoride concentrations measured in spot samples using an ion-selective electrode and adjusted for specific	Urinary fluoride concentrations were measured in spot samples using an ion-selective electrode and adjusted for specific gravity	TSH levels	Approximately 1,000, representing 6,914,124 adults in Canada aged 18–79 years

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
		members of the Canadian Armed Forces; institutionalised individuals; and those living in remote areas. Sample restricted to participants aged over 18 years and those who provided urine/tap water samples, excluding pregnant women and those with iodine levels above 2.37 µmol/L.	gravity; duration of exposure not specified			
Sleep disturbance (pineal gland functioning)						
Cunningham *et al.* (2021) (CHMS)	Cross-sectional survey Canada	Participants in the CHMS: randomly selected individuals aged 3–79 years living in Canada; excluded residents of reserves and Aboriginal settlements; full-time members of the Canadian Armed Forces; institutionalised individuals; and those living in remote areas or northern territories. Sample restricted to participants aged over 15 years and those who provided urine/tap water samples.	Fluoridated (missing or mixed fluoridation data analysed separately); duration of exposure not specified	Not fluoridated	Sleep disturbance due to reduced functioning of pineal gland	Urinary fluoride sample: 1,303; water fluoride sample: 1,016

3.6.2 Quality assessment

A summary of the quality assessment of the four papers that examined endocrine conditions, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 26 (see Appendix M for the full quality assessment). Three of the four papers received an overall rating of low quality, while one received a rating of moderate quality; this was the ecological study that found evidence for a higher prevalence of hypothyroidism in fluoridated areas of England [11]. However, it is important to note that although an ecological study can identify theoretical relationships, it cannot be used to prove or disprove causality. In addition, cross-sectional studies are useful for estimating prevalence; planning and evaluating health services; and identifying theoretical relationships, but cannot prove or disprove causality, as they collect exposure and outcome information at the same time.

Table 26 Summary of quality assessment ratings for studies examining endocrine conditions

Item	Peckham *et al.* (2015)	Barberio *et al.* (2017b) (CHMS)	Malin *et al.* (2018) (CHMS)	Cunningham *et al.* (2021) (CHMS)
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	No	No	No
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Partial	Unable to determine
Quality rating	3.0 (moderate)	2.0 (low)	2.5 (low)	2.0 (low)

3.6.3 Findings

The summary findings of the four papers that examined endocrine conditions are presented in Table 27 and Table 28 (see Appendix N for full findings).

One ecological study [11] found that general practices with CWF in England had nearly twice the odds of recording a high prevalence of hypothyroidism compared with general practices with non-fluoridated water.

Evidence from two of the papers based on the same cross-sectional study (the CHMS) was mixed. One analysis [51] found no evidence for a simple association between fluoride exposure and levels of TSH; however, a later analysis of the same dataset [53] found a moderating effect of iodine deficiency, such that for adults with moderate to severe iodine deficiency, a 1 ppm increase in specific gravity-adjusted urinary fluoride was associated with a 0.35 milli-international units per litre (mIU/L) increase in TSH; that is, iodine-deficient adults exposed to higher levels of fluoride were at higher risk for underactive thyroid gland activity (hypothyroidism).

Another paper based on the same cross-sectional study dataset [52] generally found no association between fluoride exposure (measured by tap water concentrations and specific gravity-adjusted urinary concentrations) and a range of self-reported sleep outcomes, including sleeping more (or less) than the

recommended duration, trouble sleeping, and daytime sleepiness, with one exception: for every 0.5 ppm increase in tap water fluoride concentration, there was a 34% increased relative risk of self-reporting sleeping less than the recommended duration.

Overall, although the effects of fluoride on thyroid functioning have long been observed (see Appendix A for a fuller discussion), the evidence for an association between CWF specifically and endocrine conditions and outcomes was mixed, and the findings were based on a small number of ecological and cross-sectional studies, not high-quality cohort studies. Therefore, no relationship has been firmly established. A high-quality prospective longitudinal study based on individual-level exposures and taking account of all potential confounding factors and effect modifiers (such as iodine) is required in order to strengthen the evidence base on the relationship between CWF and endocrine conditions.

Table 27 Summary of findings for studies examining endocrine conditions

Paper	Comparisons	Method of measurement	Summary of findings
Hypothyroidism prevalence			
Peckham *et al.* (2015)	Fluoridated vs. non-fluoridated regions	General practice-level hypothyroidism prevalence	General practices in fluoridated regions had nearly twice the odds of recording a high level of hypothyroidism compared with practices in non-fluoridated regions (odds ratio: 1.935 (1.388–2.699), p<0.001).
Diagnosed thyroid condition			
Barberio *et al.* (2017b) (CHMS) ●	Outcome regressed on urinary fluoride	Self-reported diagnosis of a thyroid condition: yes/no	No association between fluoride exposure measured by urinary fluoride and self-reported diagnosis of a thyroid condition
Barberio *et al.* (2017b) (CHMS) ●	Outcome regressed on fluoride concentration in tap water	Self-reported diagnosis of a thyroid condition: yes/no	No association between fluoride exposure measured by tap water and self-reported diagnosis of a thyroid condition
Low TSH levels			
Barberio *et al.* (2017b) (CHMS) ●	Outcome regressed on urinary fluoride	Low TSH levels; blood samples collected by a phlebotomist using standardised venepuncture method; quantification of TSH in serum determined using third-generation assay analyser with a chemiluminescent detection system	No association between fluoride exposure measured by urinary fluoride and low TSH levels
Barberio *et al.* (2017b) (CHMS) ●	Outcome regressed on fluoride concentration in tap water	Low TSH levels; blood samples collected by a phlebotomist using standardised venepuncture method; quantification of TSH in serum determined using third-generation assay analyser with a chemiluminescent detection system	No association between fluoride exposure measured by tap water and low TSH levels
High TSH levels			
Barberio *et al.* (2017b) (CHMS) ●	Outcome regressed on urinary fluoride	High TSH levels; blood samples collected by a phlebotomist using standardised venepuncture method; quantification of TSH in serum determined using third-generation	No association between fluoride exposure measured by urinary fluoride and high TSH levels

Paper	Comparisons	Method of measurement	Summary of findings
		assay analyser with a chemiluminescent detection system	
Barberio *et al.* (2017b) (CHMS) 🟢	Outcome regressed on fluoride concentration in tap water	High TSH levels; blood samples collected by a phlebotomist using standardised venepuncture method; quantification of TSH in serum determined using third-generation assay analyser with a chemiluminescent detection system	No association between fluoride exposure measured by tap water and high TSH levels
Less than recommended sleep duration			
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on fluoride concentration in tap water	Self-reported habitual sleep duration, reported to the closest half-hour, categorised as lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups	For every 0.5 mg/L increase in tap water fluoride concentration, there was a 34% increased relative risk of reporting sleeping less than the recommended duration.
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on urinary fluoride	Self-reported habitual sleep duration, reported to the closest half-hour, categorised as lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups	No association between change in urinary fluoride and relative risk of sleeping less than recommended duration
More than recommended sleep duration			
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on fluoride concentration in tap water	Self-reported habitual sleep duration, reported to the closest half-hour, categorised as lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups	No association between change in fluoride concentration in tap water and relative risk of sleeping more than recommended duration
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on urinary fluoride	Self-reported habitual sleep duration, reported to the closest half-hour, categorised as lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups	No association between change in urinary fluoride and relative risk of sleeping more than recommended duration
Trouble sleeping			
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on fluoride concentration in tap water	Self-reported frequency of sleep problems, single question with 5-point response scale	No association between change in fluoride concentration in tap water and relative risk of trouble sleeping
Cunningham *et al.* (2021) (CHMS) 🟢	Outcome regressed on urinary fluoride	Self-reported frequency of sleep problems, single question with 5-point response scale	No association between change in urinary fluoride and relative risk of trouble sleeping
Daytime sleepiness			

Paper	Comparisons	Method of measurement	Summary of findings
Cunningham *et al.* (2021) (CHMS) ●	Outcome regressed on fluoride concentration in tap water	Self-reported frequency of daytime sleepiness, single question with 5-point response scale	No association between change in fluoride concentration in tap water and relative risk of daytime sleepiness
Cunningham *et al.* (2021) (CHMS) ●	Outcome regressed on urinary fluoride	Self-reported frequency of daytime sleepiness, single question with 5-point response scale	No association between change in urinary fluoride and relative risk of daytime sleepiness

Table 28 Summary of findings for study examining endocrine conditions using regression analysis

Paper	Key variables	Method of measurement	Summary of findings
TSH levels			
Malin *et al.* (2018) (CHMS) ●	Specific gravity-adjusted urinary fluoride; urinary iodine; interaction betweenspecific gravity-adjusted urinary fluoride*; urinary iodine	Serum TSH levels measured using a third-generation assay analyser equipped with a chemiluminescent detection system	No evidence for an association between urinary fluoride and TSH in the absence of iodine status; however, moderate to severe iodine deficiency revealed an association of a 0.35 mIU/L increase in TSH for every 1 mg/L increase in urinary fluoride (specific gravity-adjusted) (interaction)

3.7 Renal conditions

One study [41] examined renal conditions; in this case, renal calculi (kidney stones). Most ingested fluoride is excreted via the kidneys, which are therefore exposed to high concentrations of fluoride [89].

3.7.1 Study characteristics

The summary characteristics of the single ecological study [41] that examined renal conditions are presented in Table 29 (see Appendix L for full study characteristics). This was an ecological study that monitored the health effects of water fluoridation arrangements in England. The study compared rates of selected non-dental health outcomes (in this case, renal calculi) between areas according to whether the level of fluoride in drinking water was adjusted (fluoridated) or not (non-fluoridated).

Table 29 Summary of study characteristics for study examining renal conditions

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Renal calculi (kidney stones)						
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Kidney stones	Population: fluoridated areas 37,971,918, non-fluoridated areas 274,884,530 Cases: fluoridated areas: 18,579; non-fluoridated areas: 141,963

3.7.2 Quality assessment

A summary of the quality assessment of the single ecological study [41] that examined renal calculi, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 30 (see Appendix M for the full quality assessment). The study received an overall rating of moderate. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 30 Summary of quality assessment ratings for study examining renal conditions

Item	Young *et al.* (2015)
3. Was the participation rate of eligible persons at least 50%?	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited
Quality rating	3.0 (moderate)

3.7.3 Findings

The summary findings of the single ecological study [41] that examined renal calculi are presented in Table 31 (see Appendix N for full findings). The study found that the incidence of renal calculi (kidney stones) was lower in fluoridated areas than in non-fluoridated areas, controlling for age, gender, deprivation, and ethnicity.

Table 31 Summary of findings for study examining renal conditions

Paper	Comparisons	Method of measurement	Summary of findings
Renal calculi (kidney stones)			
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	Incidence of kidney stones inpatient consultant episodes per lower super output area in England recorded in hospital episode statistics; admission dates between April 2007 and March 2013	Incidence of kidney stones lower in fluoridated areas (incidence rate ratio: 7.9 (9.6–6.2), p<0.001)

3.8 Birth or birthing abnormalities

Three papers [41,69,74] examined birth or birthing abnormalities; the incidence of Down syndrome was the outcome of interest in two of these studies, and the incidences of trisomies, stillbirths, neural tube defects, clefts, and preterm births were each examined by one study. A cytogenetic effect on the developing foetus is theoretically plausible, as fluoride can cross the placenta (Appendix A). Birth defects can include many different congenital abnormalities. In relation to water fluoridation, the most concern has been expressed about the possible association of fluoridation and the occurrence of Down syndrome in babies born to women exposed to fluoride. However, no statistically significant association has been found in previous reviews [5]. It has been proposed that the bacteria that cause dental caries and periodontitis can enter the bloodstream and trigger an inflammatory cascade, leading to preterm birth; fluoride, in reducing caries, can thus theoretically lower the risk of preterm birth [74]. However, there is limited evidence to support an association between periodontal disease and preterm birth, which have many more substantial risk factors in common, including smoking [90,91].

3.8.1 Study characteristics

The summary characteristics of the three ecological studies that examined birth or birthing abnormalities are presented in Table 32 (see Appendix L for full study characteristics). Two of these studies [41,69] compared the incidence of outcomes of interest between fluoridated and non-fluoridated areas; the third study [74] examined the association between maternal dental cleaning and exposure to fluoridated water during pregnancy and the incidence of preterm birth.

Table 32 Summary of study characteristics for studies examining birth or birthing abnormalities

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
Lowry *et al.* (2003)	Ecological or correlational study England	Mothers and their newborn children in northern England	Artificially fluoridated areas in northern region of England (>0.9 ppm); duration of exposure not specified	Non-fluoridated areas in northern region of England (<0.3 ppm)	Stillbirths and congenital abnormalities (all trisomies, Down syndrome, neural tube defects, and clefts)	Not reported
Young *et al.* (2015)	Ecological or correlational study England	Mothers and their newborn children in England	Fluoridated areas, aims to fluoridate to 1 ppm; duration of exposure not specified	Non-fluoridated areas	Down syndrome	Population (live births): fluoridated areas 303,818, non-fluoridated areas 2,423,482 Cases: fluoridated areas 658; non-fluoridated areas 5,961
Zhang *et al.* (2019)	Ecological or correlational study USA	Mothers and their newborn children in Massachusetts, USA	Exposure to CWF (with or without dental cleaning); duration of exposure not specified	Not exposed to CWF (with or without dental cleaning)	Preterm birth or birth prior to 37 weeks' completed gestation	9,234

3.8.2 Quality assessment

A summary of the quality assessment of the three ecological studies [41,69,74] that examined birth and birthing abnormalities, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 33 (see Appendix M for the full quality assessment). Two of these studies received an overall rating of moderate quality, while the third received a rating of low quality. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 33 Summary of quality assessment ratings for studies examining birth and birthing abnormalities

Item	Lowry *et al.* (2003)	Young *et al.* (2015)	Zhang *et al.* (2019)
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	No
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Partial
Quality rating	3.0 (moderate)	3.0 (moderate)	2.5 (low)

3.8.3 Findings

A summary of the findings of the three studies that examined birth and birthing abnormalities is presented in Table 34 (see Appendix N for full findings). There was no association between exposure to fluoridated water and the incidence of Down syndrome[41,69], trisomies, neural tube defects, clefts, or stillbirths [69]. One study [74] found that women who received dental cleaning and were exposed to artificially fluoridated water, along with those who received dental cleaning alone, had a significantly lower incidence of preterm births compared with those who had neither, after controlling for confounding variables. This study found that CWF alone had no association with the incidence of preterm births.

Table 34 Summary of findings for studies examining birth and birthing abnormalities

Paper	Comparisons	Method of measurement	Summary of findings
Down syndrome			
Lowry *et al.* (2003)	Fluoridated vs. non-fluoridated regions	Cases of Down syndrome, all cases of a congenital abnormality with a final postnatal diagnosis of trisomy 21 using International Classification of Diseases, 9th Revision (ICD-9) codes	No significant association between water fluoridation and incidence of Down syndrome
Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	Cases of Down syndrome according to lower-tier local authority, recorded in the National Down Syndrome Cytogenetic Register, including live births, stillbirths, late miscarriages, and terminations of	No evidence of a difference in rate of Down syndrome in fluoridated and non-fluoridated areas when adjusted for maternal age

Paper	Comparisons	Method of measurement	Summary of findings
		pregnancy with foetal anomaly, from 2009 to 2012 inclusive	
All trisomies			
Lowry *et al.* (2003)	Fluoridated vs. non-fluoridated regions	All cases of a congenital abnormality with a final postnatal diagnosis of a trisomy (trisomies 21, 13, and 18 only, ICD-9 codes 758.0, 758.1, and 758.2)	No significant association between water fluoridation and incidence of trisomies
Neural tube defects			
Lowry *et al.* (2003)	Fluoridated vs. non-fluoridated regions	All cases of a congenital abnormality with a final postnatal diagnosis of a neural tube defect (as defined by the EUROCAT (European Registration of Congenital Anomalies and Twins) system of classification, ICD-9 codes 740.0, 740.1, 740.2, 741.0, 741.9, and 742.0)	No significant association between water fluoridation and incidence of neural tube defects
Clefts			
Lowry *et al.* (2003)	Fluoridated vs. non-fluoridated regions	All cases of a congenital abnormality with a final postnatal diagnosis of a facial cleft (cleft palate, cleft lip with or without cleft palate, Pierre Robin syndrome, ICD-9 codes 749.0, 749.1, 749.2, and 756.03)	No significant association between water fluoridation and incidence of clefts
Stillbirths			
Lowry *et al.* (2003)	Fluoridated vs. non-fluoridated regions	All stillbirths between 1 January 1989 and 31 December 1998 identified from the Northern Perinatal Mortality Survey	No significant association between water fluoridation and incidence of stillbirth
Preterm births			
Zhang *et al.* (2019)	Neither dental cleaning nor CWF vs. dental cleaning only vs. CWF only vs. dental cleaning and CWF	Incidence of preterm births (<37 weeks)	Women with dental cleaning alone and dental cleaning plus CWF have a significantly lower incidence of preterm births compared with those who had neither, after controlling for confounding factors (adjusted risk ratio: 0.74 (0.55–0.98) for dental cleaning alone, 0.74 (0.57–0.95) for dental cleaning and CWF). No impact of CWF alone on preterm births.

3.9 Infant abnormalities

One case-control study with ecological assignment of CWF status [64] examined infant abnormalities; in this case, sudden infant death syndrome (SIDS) was the outcome of interest. The biological mechanism by which fluoride is linked to SIDS is not clear and the study did not present a rationale for the proposed association.

3.9.1 Study characteristics

The summary characteristics of the single ecological study [64] that examined infant abnormalities are presented in Table 35 (see Appendix L for full study characteristics). This study used data from the New Zealand Cot Death Study to determine whether exposure to fluoridated water supplies prenatally (in

drinking water) or postnatally at the time of death (through feeding with breast milk or formula) was associated with SIDS.

Table 35 Summary of study characteristics for study examining infant abnormalities

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
SIDS						
Dick *et al.* (1999)	Case-control study with ecological assignment of CWF status New Zealand	New Zealand Cot Death study covered 78% of births in New Zealand between November 1987 and October 1990	Fluoridated areas (designated >80% (average pre- and postnatal exposure) of residences fluoridated); target level 1.0 ppm; duration of exposure up to 1 year; partially fluoridated areas (20–80% of residences fluoridated) excluded from analysis	Non-fluoridated areas (designated <20% of residences fluoridated (average pre- and postnatal exposure))	Cases of SIDS	Total: 2,285 (485 SIDS cases; 1,800 controls)

3.9.2 Quality assessment

A summary of the quality assessment of the single case-control study with ecological assignment of CWF status [64] that examined infant abnormalities, using the NHLBI quality assessment tool for case-control studies, is reported in Table 36 (see Appendix M for the full quality assessment). The study received an overall rating of low. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 36 Summary of quality assessment rating for study examining infant abnormalities

Item	Dick *et al.* (1999)
4. Did the authors include a sample size justification?	No
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	Partly
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Limited
Quality rating	2.5 (low)

3.9.3 Findings

A summary of the findings of the single study that examined infant abnormalities (in this case, SIDS) is presented in Table 37 (see Appendix N for full findings). The study found no association between SIDS and prenatal exposure to fluoridated water [64]. Postnatally, the study also examined the association between SIDS and water fluoridation status and feeding method (breastfeeding compared with formula feeding). No higher risk of SIDS was associated with either breastfeeding or formula feeding in fluoridated areas compared with non-fluoridated areas, nor was there any evidence of an interaction between water fluoridation status and feeding.

Table 37 Summary of findings for study examining infant abnormalities

Paper	Comparisons	Method of measurement	Summary of findings
SIDS			
Dick *et al.* (1999)	New Zealand fluoridated vs. non-fluoridated regions; North Island fluoridated vs. non-fluoridated regions	Cases of SIDS drawn from the New Zealand Cot Death Study, 1987–1990	Prenatal exposure to fluoridated water was not associated with a higher risk of SIDS in either the full New Zealand sample or the North Island sample (which excluded the South Island, where fluoride supplementation was more prevalent at the time of the study).
	Breastfed infants in fluoridated vs. non-fluoridated regions vs. formula-fed infants in fluoridated vs. non-fluoridated regions	Cases of SIDS drawn from the New Zealand Cot Death Study, 1987–1990	No higher risk of SIDS associated with breastfeeding in fluoridated areas; no evidence for interaction between fluoridated water supplies and infant feeding
	Formula-fed infants in fluoridated vs. non-fluoridated regions	Cases of SIDS drawn from the New Zealand Cot Death Study, 1987–1990	Fluoridated formula feeding showed no higher risk of SIDS compared with non-fluoridated formula feeding.

3.10 All-cause mortality

One study [41] examined all-cause mortality. No direct link between fluoride and excess death is proposed; rather, the hypothesis here is that the summative negative health effects of CWF contribute to excess deaths.

3.10.1 Study characteristics

The summary characteristics of the single ecological study [41] that examined all-cause mortality are presented in Table 38 (see Appendix L for full study characteristics). This was an ecological study that monitored the health effects of water fluoridation arrangements in England. The study compared rates of selected non-dental health outcomes (in this case, deaths) between areas, according to whether the drinking water was fluoridated or not.

Table 38 Summary of study characteristics for study examining all-cause mortality

Paper	Study design Country	Population	Exposure	Comparator	Outcome	Sample size for analysis
All-cause mortality						
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Count of deaths	Population: fluoridated areas 25,314,612, non-fluoridated areas 183,256,350 Cases: fluoridated areas: 233,922; non-fluoridated areas: 1,602,206

3.10.2 Quality assessment

A summary of the quality assessment of the single ecological study [41] that examined all-cause mortality, using the NHLBI quality assessment tool for observational cohort and cross-sectional studies, is reported in Table 39 (see Appendix M for the full quality assessment). The study received an overall rating of moderate. However, it is important to note that an ecological study can identify theoretical relationships but cannot be used to prove or disprove causality.

Table 39 Summary of quality assessment rating for study examining all-cause mortality

Item	Young *et al.* (2015)
3. Was the participation rate of eligible persons at least 50%?	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes
13. Was loss to follow-up after baseline 20% or less?	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Unable to determine
Quality rating	3.0 (moderate)

3.10.3 Findings

A summary of the findings of the single study that examined all-cause mortality is presented in Table 40 (see Appendix N for full findings). The study found some evidence that the death rate from all recorded causes was lower in fluoridated areas than in non-fluoridated areas, but the effect size was small [41].

Table 40 Summary of findings for study examining all-cause mortality

Paper	Comparisons	Method of measurement	Summary of findings
All-cause mortality			

Young *et al.* (2015)	Fluoridated vs. non-fluoridated regions	Count of deaths, obtained at lower super output area level from the Office of National Statistics in England from January 2009 to January 2012	Some evidence that rate of deaths from all recorded causes was lower in fluoridated areas than non-fluoridated areas (incidence rate ratio: 1.3 (−2.5 to −0.1), p=0.04), but effect size was small

3.11 Grading of Recommendations, Assessment, Development and Evaluations rating

The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach to rating quality of evidence and strength of recommendations was applied to the included papers as outlined in Section 2.7.3. Scores for each of the eight evaluated outcomes are displayed in Table 41. It is not appropriate to apply GRADE to ecological studies, which cannot contribute evidence for causality and are therefore not included in the hierarchy of evidence (see Section 2.5). Therefore, ecological studies and other study types are addressed separately in the assessment.

The a priori and final rating was “very low” for all outcomes for which only ecological studies were available; as such studies cannot contribute evidence for causality, they can offer very low levels of certainty about the effects they describe. For all other outcomes, the a priori rating was “low”, because only observational studies were available in each case. In the case of sleep disturbance, the a priori rating was “very low” because evidence from only one observational study was available; the full GRADE approach was not applied to this outcome, as one study does not constitute a body of evidence.

In the absence of randomised studies to which risk of bias assessments might have been applied, we used the quality assessment scores to make determinations about downgrading based on risk of bias. Indirectness was deemed not to be applicable in all cases.

All outcomes received a final rating of “very low” certainty in the evidence. In most cases, the rating was downgraded due to the low quality of individual studies (i.e. risk of bias), inconsistent findings across studies, and imprecision (generally marked by the presence of wide confidence intervals in many or all studies).

It is important to note that using the GRADE system, observational studies (such as the cross-sectional studies and cohort studies included in this review), even those with very good design and implementation, generally provide only a moderate degree of evidence. Only randomised controlled trials provide a high degree of certainty; however, it is arguably both infeasible and unethical to attempt to investigate CWF using a randomised controlled trial. Thus, under GRADE, policy and regulatory decisions will have to be made based on observational studies or moderate-quality evidence at best.

Table 41 GRADE rating for all outcomes

Outcome	A priori rating	Downgrade for					Upgrade for			Final rating
		Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Large consistent effect	Dose response	Confounders only reducing size of effect	
Bone characteristics (cross-sectional studies)	Low	No limitation	Serious limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low
Bone characteristics (ecological studies)	Very low	Not applied – ecological studies only								Very low
Fractures (cross-sectional studies)	Low	No limitation	No limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low
Fractures (ecological studies)	Very low	Not applied – ecological studies only								Very low
IQ/ neuropsychological development	Low	Serious limitation	Serious limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low
Neurodevelopmental disorders	Low	Serious limitation	Serious limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low
Bone cancers (case-control studies)	Low	Serious limitation	Serious limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low
Bone cancers (ecological studies)	Very low	Not applied – ecological studies only								Very low
Other cancers	Very low	Not applied – ecological studies only								Very low
Thyroid disorders (cross-sectional studies)	Low	Serious limitation	Serious limitation	Not applicable	Serious limitation	No limitation	No upgrade	Not applicable	No upgrade	Very low

Outcome	A priori rating	Downgrade for	Upgrade for	Final rating
Thyroid disorders (ecological studies)	Very low	Not applied – single ecological study		Very low
Sleep disturbance	Low	Not applied – single study		Very low
Renal conditions (renal calculi)	Very low	Not applied – single ecological study		Very low
Down syndrome	Very low	Not applied – ecological studies only		Very low
Congenital abnormalities	Very low	Not applied – single ecological study		Very low
Stillbirths	Very low	Not applied – single ecological study		Very low
Preterm births	Very low	Not applied – single ecological study		Very low
Infant abnormalities (SIDS)	Very low	Not applied – single case-control study		Very low
All-cause mortality	Very low	Not applied – single ecological study		Very low

4 Discussion

4.1 Summary of findings

A total of 30 studies reported across 37 papers meeting the review eligibility criteria were identified from all stages of the search process. Many studies examined multiple outcomes.

4.1.1 Bone health

4.1.1.1 Bone characteristics

Eight papers [31,32,45,46,57,60,68,73] examined the association between fluoridated water and bone characteristics. Four of the eight papers presented data from three cross-sectional surveys, three papers were ecological studies, and one paper was a cross-sectional study that allocated exposure on an ecological or population basis, which should be treated as an ecological study for the purposes of contributing evidence for causality. Overall, the evidence from these studies for the influence of CWF on BMD was mixed. Although a number of studies found associations between higher BMD and exposure to water fluoridation in certain skeletal areas (such as the lumbar spine), contradictory findings also exist, and a large number of analyses found no association. Therefore, no theoretical relationship has been firmly established. Additionally, one ecological study found no association between the incidence of osteoporosis and water fluoridation status [68].

4.1.1.2 Fractures

Ten papers [30–32,41,45,46,63,66–68] presented data from 9 studies that examined the association between water fluoridation status and the incidence of a range of fractures, most commonly hip fractures. Seven of the 10 papers were based on ecological studies and 3 papers were based on 2 cross-sectional surveys. Overall, the evidence from these papers for an association between CWF and fracture incidence was mixed, with most analyses pointing to a neutral or, in a few analyses, possible protective effect of fluoridation, although only hip fracture has been extensively studied. It is important to note that none of the included analyses controlled for osteoporosis, which is the leading risk factor for hip fracture.

4.1.2 Neuropsychological outcomes

Seven papers [38,49,50,54–56,59] based on four studies examined the association between fluoridated water and neuropsychological outcomes. Four papers [54–56,59] based on two studies examined IQ in childhood and adulthood, and one additional paper examined aspects of neuropsychological development in infancy and childhood, which conceptually maps closely to IQ [38].

The studies investigating the influence of fluoride on IQ and neuropsychological development have mixed findings, variously reporting null, positive, and inverse associations between fluoride exposure and IQ in childhood (however, the small number of positive associations should not be interpreted as evidence for a beneficial effect of fluoride on IQ). Two of the three cohort studies had high loss to follow-up (the MIREC study [54–56] and Ibarluzea *et al.* [38]) and the MIREC longitudinal cohort study has methodological issues that call into question the validity of the findings presented in these papers [54–56].

The remaining two papers [49,50], based on one cross-sectional study (CHMS), present conflicting findings with respect to diagnosis of ADHD, and one analysis demonstrated stronger associations between fluoride exposure and hyperactivity for older youth. A high-quality prospective longitudinal study based on individual-level exposures, taking account of all potential confounding factors, effect modifiers, and

cluster design effect, is required in order to strengthen the evidence base on neuropsychological outcomes.

4.1.3 Cancer

4.1.3.1 Bone cancers

Eleven papers [20,41,47,48,58,61,65,68,70–72] presented data from 10 studies examining the association between bone cancers and fluoridated water status. Eight were ecological papers and three were case-control papers. Osteosarcoma was the cancer examined in 10 papers (9 studies), bone cancers in general were examined by 2 papers, and Ewing sarcoma was examined by 1 paper. The evidence from the ecological studies does not suggest any association between CWF and the incidence of bone cancers. In addition, the relationship between a diagnosis of osteosarcoma and exposure to artificially fluoridated water is unlikely, based on the evidence from case-control studies. Therefore, no relationship can be established.

4.1.3.2 Other cancers

Two ecological studies examined other cancers. One ecological study examined differences in incidences of any cancer and bladder cancer between fluoridated and non-fluoridated areas [41], while the other ecological study examined differences in secondary bone cancer incidences according to the extent of fluoridation implementation [62]. The first, an ecological study conducted in England, found a lower incidence of bladder cancer and all cancers in fluoridated areas (compared with non-fluoridated areas) after adjusting for confounders. While these results suggest some possible protective effects of fluoridation against some forms of cancer, that association is based on an ecological study design and cannot be considered causal. In the second ecological study of cancer patients, no relationship was found between the percentage of county-level access to fluoridated water and the prevalence of secondary bone cancer.

4.1.4 Endocrine conditions

Four papers [11,51–53], presenting data from two studies, examined the association between f artificially fluoridated water and endocrine conditions. Three papers [11,51,53], based on two studies, examined a range of outcomes related to thyroid functioning, including incidence of diagnoses of thyroid disorders and TSH levels, while one paper [52] examined the incidence of sleep disturbances, which the study authors attribute to the functioning of the pineal gland.

Overall, although the effects of fluoride on thyroid functioning have long been observed (see Appendix A for a fuller discussion), the evidence for an association between CWF specifically and thyroid conditions and outcomes was mixed, and the findings were based on a small number of ecological and cross-sectional studies, not high-quality cohort studies. Therefore, no relationship has been firmly established. A high-quality prospective longitudinal study based on individual-level exposures and taking account of all potential confounding factors and effect modifiers (such as iodine) is required in order to strengthen the evidence base on the relationship between CWF and thyroid and other endocrine conditions. One cross-sectional survey paper [52] generally found no association between fluoride exposure (measured by tap water concentrations and specific gravity-adjusted urinary concentrations) and a range of self-reported sleep outcomes, including sleeping more than the recommended duration, trouble sleeping, and daytime sleepiness, although it found some evidence for a higher risk of sleeping less than the recommended amount with higher fluoride exposure.

4.1.5 Renal conditions

One ecological study [41] monitored the health effects of water fluoridation arrangements in England. The study compared rates of selected non-dental health outcomes (in this case, renal calculi) between areas according to whether the level of fluoride in drinking water was adjusted (fluoridated) or not (non-fluoridated). The study found that the incidence of renal calculi (kidney stones) was lower in fluoridated areas than in non-fluoridated areas, controlling for age, gender, deprivation, and ethnicity.

4.1.6 Birth or birthing abnormalities

Three ecological studies [41,69,74] examined birth or birthing abnormalities; the incidence of Down syndrome was the outcome of interest in two of these studies, and the incidences of trisomies, stillbirths, neural tube defects, clefts, and preterm births were each examined by one study. The studies found no association between exposure to fluoridated water and the incidence of Down syndrome, trisomies, neural tube defects, clefts, or stillbirths. One study [74] found that women who received dental cleaning and were exposed to artificially fluoridated water, along with those who received dental cleaning alone, had a significantly lower incidence of preterm births compared with those who had neither, after controlling for confounding variables. This study found that CWF alone had no association with the incidence of preterm births.

4.1.7 Infant abnormalities

One case-control study with ecological assignment of CWF status [64] examined infant abnormalities; in this case, SIDS was the outcome of interest. The study found no association between SIDS and prenatal exposure to fluoridated water. Postnatally, the study also examined the association between SIDS and water fluoridation status and feeding method (breastfeeding compared with formula feeding). No higher risk of SIDS was associated with either breastfeeding or formula feeding in fluoridated areas compared with non-fluoridated areas, nor was there any evidence of an interaction between water fluoridation status and feeding.

4.1.8 All-cause mortality

One ecological study [41] examined all-cause mortality. The study found some evidence that the death rate from all recorded causes was lower in fluoridated areas than in non-fluoridated areas, but the effect size was small.

4.1.9 Summary statement

There is very low-quality evidence (see Section 3.11) for the impact of artificially fluoridated water on systemic health, and there continues to be no definitive evidence that artificially fluoridated water is associated with negative health outcomes. The epidemiological study designs employed, except for some of those used for neuropsychological outcomes and bone cancers, are generally correlational or cross-sectional methods used for generating rather than testing theories and cannot be used to robustly establish causality. High-quality prospective longitudinal studies or case-control studies (for rare diseases) based on individual-level exposures, taking account of all potential confounding factors and effect modifiers, are required in order to strengthen the evidence base on bone mineral density, osteoporosis, fracture, neuropsychological outcomes, cancer, endocrine conditions, renal conditions, birth and birthing abnormalities, infant abnormalities, and all-cause mortality. It would be an advantage if these were conducted across the many countries (including Ireland) that use CWF as a public health intervention to prevent caries.

We found no studies meeting the inclusion criteria that examined the relationship between artificially fluoridated water and diabetes [92] or eye melanoma [93], although we did find one study on each topic that covered the topics in optimal fluoridation (mix of natural and artificial fluoridation) areas in the USA. We also did not find studies examining the relationship between artificially fluoridated water and cardiovascular or immune outcomes.

4.2 Comparison with other reviews

4.2.1 Bone health

The authors of the York review [5], published in 2000, concluded that water fluoridation at levels aimed at preventing dental caries had little effect on fracture risk, either protective or deleterious. There were no definite patterns of association for any of the fractures studied [5]. The Australian National Health and Medical Research Council (NHMRC) review, published in 2007 [22], supported the conclusion of the York review on bone fractures. A review of the scientific evidence on fluoridation on behalf of the Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor [94], published in 2014, concluded that, based on the available evidence, there was no appreciable risk of bone fractures arising from artificially fluoridated water. In the 2021 update of that review, the conclusion remained unchanged [95]. The NHMRC updated its review in 2017 [96] and found that there was reliable evidence that fluoridation at Australia's current levels was not associated with hip fracture. In 2019, a Canadian Agency for Drugs and Technologies in Health (CADTH) report concluded that there was consistent evidence for no association between water fluoridation and hip fracture, and that there was insufficient evidence for an association between water fluoridation at Canada's current levels and osteoporosis [6].

4.2.2 Neuropsychological outcomes

The 2000 York review and the 2007 Australian NHMRC review did not specifically examine the issue of fluoridated water and IQ [5,22]. The scientific evidence on fluoridation published by the Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor [94] in 2014 reported on the study by Broadbent *et al.* [59] in New Zealand (published online in May 2014 and later in print in 2015), which revealed no evidence that exposure to fluoridated water in New Zealand affects neuropsychological development or IQ. The 2014 report's authors concluded that, based on the available evidence, there was no appreciable effect on cognition arising from artificially fluoridated water. In its updated review in 2017 [96], the NHMRC found that there was reliable evidence that CWF at Australia's current levels was not associated with cognitive dysfunction or lowered intelligence. In 2019, CADTH [6] concluded that there was limited evidence for no association between water fluoridation at Canada's current levels and IQ or cognitive function. Following the 2021 update of its report, the Office of the Prime Minister's Chief Science Advisor in New Zealand reported that there was no convincing evidence of neuropsychological effects at fluoride concentrations achieved by artificial fluoridation of water supplies in New Zealand [95]. In 2019, the National Toxicology Program in the USA published their draft monograph on a systematic review of the neurodevelopment and cognitive health effects of fluoride exposure. While still in draft form, the review is broadly in accordance with the findings of this review, finding inconsistent evidence for an effect on neurodevelopment where fluoride concentrations are below 1.5 ppm [84]. A 2021 review by Miranda *et al.* [83] of studies examining the association between fluoride exposure and IQ found evidence for an association between high levels of fluoride in drinking water (>2.0 mg/L) and IQ deficits. The Miranda *et al.* review found similar methodological problems as this review, including high heterogeneity, a preponderance of cross-sectional studies, and very low-quality evidence overall.

4.2.3 Cancer

4.2.3.1 Bone cancer

The authors of the York review [5] examined 11 studies relating to bone cancer and osteosarcoma and concluded that no clear association exists between water fluoridation and overall cancer incidence or mortality (for 'all-cause' cancer and specifically for bone cancer and osteosarcoma).

Australia's NHMRC, in its 2007 review [22], found four studies in addition to those identified in the York review that examined the association between cancer and water fluoridation. Only one of these studies compared the fluoride exposure of histologically confirmed osteosarcoma cases with matched controls. The authors of the NHMRC review stated that after adjusting for significant differences at baseline between the cases and controls, the results presented in the Bassin *et al.* (2006) paper, based on data from the Harvard Fluoride Osteosarcoma Study, [47] suggested a higher risk of osteosarcoma among young males (but not young females) who were exposed to artificial water fluoridation.

However, three later analyses of the data from the Harvard study were unable to replicate this finding. First, the NHMRC review drew attention to a letter by Douglass and Joshipura [97] (co-investigators on the Harvard Fluoride Osteosarcoma Study) to the editor of *Cancer Causes & Control.* The letter stated that they were unable to replicate the findings of Bassin *et al.* in the broader data from the Harvard study, which included prospective cases from the same 11 hospitals as were included in the Bassin *et al.* analysis. Second, analysis of the bone samples from the Harvard study demonstrated no association between the level of fluoride content in the bone samples and risk of osteosarcoma [98]. Third, the paper by Kim *et al.* (2020) [48] included in this review presented data from the full Harvard study with a larger sample than Bassin *et al.*; this analysis failed to replicate the findings of the Bassin *et al.* and supports the assertions of Douglass and Joshipura. Therefore, despite the concerns raised initially by Bassin *et al.* in 2006, the later phases of the Harvard study did not confirm an association between fluoride and osteosarcoma risk.

The Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor's 2014 report [94] reviewed the evidence on water fluoridation and cancer. The report authors refer to the fact that almost all epidemiological studies found no association between fluoride and cancer, even after decades of exposure in some populations. They also note that although the Bassin *et al.* [47] study claimed a higher risk of osteosarcoma in young males, extensive reviews of these and other data have not found an association between exposure to fluoridated water and risk of osteosarcoma. The authors of the New Zealand report concluded that, based on the available evidence, there was no appreciable risk of cancer arising from CWF. In their 2021 update, they noted that their conclusion was unchanged [95]. The NHMRC's updated review, published in 2017 [96], found that there was reliable evidence that CWF at Australia's current levels was not associated with cancer. In its 2019 review [6], CADTH concluded that there was consistent evidence for no association between water fluoridation at Canada's current levels and the incidence of bone cancer.

4.2.3.2 Other cancers

The findings of the 2000 York review [5] on cancer studies were mixed, with small variations on either side of no effect. Individual cancers examined were bone cancers and thyroid cancer, where, once again, no clear pattern of association was seen. Overall, from the research evidence presented, the York review concluded that no association was detected between water fluoridation and mortality from any cancer, or from bone or thyroid cancers specifically. The Australian NHMRC's [22] literature review in 2007 identified three additional studies that investigated the relationship between water fluoridation and cancer incidence or mortality. Again, the results of these studies showed a mixed pattern. The NHMRC's updated review in 2017 [96] found that there was reliable evidence that CWF at Australia's current levels was not

associated with cancer In 2019, CADTH [6] concluded that there was consistent evidence for no association between water fluoridation at Canada's current levels and the overall incidence of cancer or cancer-related mortality.

4.2.4 Endocrine conditions

Goitre (or enlarged thyroid) was considered by the York review [5] team, but it only found three studies relating to this disorder. Of these, one found a statistically significant association with water fluoride level, and the other two did not find this association. Therefore, findings on goitre were mixed and, once again, no clear pattern of association was seen. The authors of the York review concluded that, overall, there was no association detected between water fluoridation and the prevalence of goitre. The Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor's [94] 2014 report stated that several other alleged effects of artificially fluoridated water on health outcomes were reviewed, including effects on endocrine function, and noted that the most reliable and valid evidence to date indicated that fluoride at levels used for CWF does not pose appreciable risks of harm to endocrine function. In 2019, CADTH [6] concluded that there was insufficient evidence for an association between water fluoridation at Canada's current levels and thyroid function or insomnia. In a 2021 update of its 2014 report, the Office of the Prime Minister's Chief Science Advisor in New Zealand reported that recent studies did not provide consistent evidence for an impact of artificially fluoridated water on thyroid hormone levels but indicated that this is an area of research that should continue to be monitored, particularly with regard to the interaction of fluoride and iodine levels [95].

4.2.5 Renal conditions

The Royal Society of New Zealand and the Office of the Prime Minister's Chief Science Advisor's [94] 2014 report stated that evidence to date indicated that fluoride concentrations at levels used for CWF did not pose appreciable risks of harm to renal function. In the 2021 update of the review, this conclusion remained unchanged [95]. In 2019, CADTH [6] concluded that there was limited evidence for an inverse association between water fluoridation at Canada's current levels and the incidence of kidney stones.

4.2.6 Birth or birthing abnormalities

In 2000, the York review team [5] reviewed six studies that examined the association between Down syndrome and water fluoridation level. Three of the six studies found a negative direction of association, one found a positive direction of association (but it was not statistically significant), one found no association, and the sixth found a non-significant positive direction of association for one set of data and a non-significant negative direction of association for the other. Australia's 2007 NHMRC review [22] reported no difference in congenital abnormalities in fluoridated and non-fluoridated areas (except for clefts, for which the risk was statistically significantly lower in the fluoridated areas). The NHMRC's updated review in 2017 [96] found that there was reliable evidence that CWF at Australia's current levels was not associated with Down syndrome. CADTH [6] concluded in its 2019 report that there was limited evidence for no association between water fluoridation at Canada's current levels and Down syndrome.

4.2.7 Infant abnormalities

The NHMRC review [22] in 2007 reported no difference in stillbirths between fluoridated and non-fluoridated areas.

4.2.8 All-cause mortality

In the 2000 York review [5], five studies examined the association between all-cause mortality and water fluoride exposure. Three of the five studies found the direction of association of water fluoridation and

mortality to be negative (i.e. there were more deaths), one found the direction of association to be positive (fewer deaths), and one found no association. No measures of the statistical significance of these associations were provided. However, for two of the studies that found a negative direction of association, the point estimate was 1.01, which was unlikely to have reflected a statistically significant effect. The Australian NHMRC review from 2007 [22] found no studies in addition to those identified in the York review examining all-cause mortality and fluoride. In 2019, CADTH [6] concluded that there was insufficient evidence for an association between water fluoridation at Canada's current levels and all-cause mortality.

4.2.9 Summary statement

The conclusions of this 2022 HRB review on each outcome of interest are in line with international evidence. The evidence on all reported outcomes will require continued surveillance, but special attention should be paid to monitoring neuropsychological and endocrine outcomes, which have theoretical plausibility but inconsistent evidence.

4.3 Excluded studies of naturally fluoridated water

Our reasons for excluding studies of areas with naturally fluoridated water within WHO guideline limits are presented in Section 2.2. While the decision to exclude these studies has led to the loss of some data on certain outcomes (e.g. diabetes), it has allowed for a much more specific and more appropriate analysis with a tightly defined exposure that has direct relevance to policy decisions being made in Ireland. Furthermore, based on the screening process from this review (see Appendix C for a list of studies excluded on full text screening) and the findings of the HRB's 2015 review [25] – which did analyse studies of areas with natural fluoridation within recommended levels alongside studies of CWF – we believed that the studies excluded on this basis would not substantially alter the conclusions of the review on artificial water fluoridation.

However, we acknowledge that the findings from studies of natural fluoridation within the WHO guideline limits may offer additional useful data for understanding the impact of fluoride. Therefore, Appendix O presents a brief overview of 20 additional studies that were excluded from the synthesis due to the presence of natural fluoride within recommended levels in the analysis, based on brief, high-level extraction of data.

In brief, the findings from the excluded studies of natural fluoride broadly mirror the findings from the main synthesis. Studies of fluoridation and bone health generally returned null or weak findings, as did studies of cancer and cardiovascular outcomes. The findings around neuropsychological and endocrine outcomes are also mirrored in the studies of natural fluoride; studies identified associations between fluoride exposure and endocrine outcomes for adolescents, diabetes, and kidney and liver function, and, at the upper end of guideline levels (1.5 ppm), associated with cognitive performance in children. However, only high-level extraction was carried out on these studies and no quality assessment has been conducted to confirm the reliability of these findings. Our conclusions from the main synthesis stand; further high-quality research is now needed in order to shed light on the impact, if any, of artificial and natural water fluoridation on these aspects of systemic health.

Also of note is one ecological study in the USA [92], which found that fluoride artificially added to drinking water to achieve optimal levels (0.7-1.2 ppm) was associated with higher incidence and prevalence of diabetes, when accounting for per capita consumption of tap water and other confounders. However, of the three fluoridation chemicals examined as part of the study (sodium fluoride, fluorosilicic acid, or sodium fluorosilicate), fluorosilicic acid was associated with lowest incidence and prevalence of diabetes; this is the chemical used in Ireland for CWF. This is an area requiring further study.

4.4 Strengths and limitations

The HRB's 2015 review [25] included studies of areas with natural fluoridation within recommended levels; these studies are excluded from this updated report. The main reason for this is that areas with concentration levels far exceeding the 1.5 ppm WHO guideline level may also have other toxic materials (e.g. heavy metals such as arsenic and lead) in the water, and so the effect of naturally occurring fluoride cannot be assessed in isolation. In addition, there is a well-accepted dose response between fluoride intake and the likelihood of developing fluorosis. Finally, exposure to fluoride at levels above 1.5 ppm is not a useful reference point for policy decisions being taken in Ireland. While this has led to the loss of some data on certain outcomes (e.g. diabetes), it has allowed for a much more specific and more appropriate analysis with a tightly defined exposure.

The approach to this search was carefully considered, with the aim of capturing all relevant evidence that would best answer the review question. The review team consulted with the information specialist who worked on the 2015 review; this was very useful in teasing out basic concepts and learning from that earlier process. The primary strength of this search is that it is an expert peer-reviewed comprehensive search, conducted across a range of reputable databases and sources and using best practice in methodology, all of which fortifies the validity of the search result. Staging the search to meet the process of the review – scoping, the main database search, the supplementary and grey literature search, the reference and citation chasing search, and the final date-specific database search – gave a thorough overview of available evidence.

Regarding the limitations of the search, only English-language reviews were considered for full-text inclusion. As a topic with confounding language (e.g. multiple types of fluorine/fluoride), use of a simple translator (e.g. Google Translate) risked mistranslating technical phrases and details. Time would not allow recruitment of a professional translator.

The evidence on the health outcomes examined is generally of very low quality (see Section 3.11) due to the study designs employed or the methodological limitations of the implementation of the study designs. The predominant study designs were ecological and cross-sectional studies, which are not adequate for inferring causality. Ireland contributed 2 of the 37 papers to the evidence base, and both were ecological or correlational studies. There is an overlap of evidence, with some data being included in multiple different papers. A number of studies or papers are linked, either because they analyse data from the same dataset or because they expand on or update the same earlier work. These linked studies are highlighted wherever they arise in the sections reporting on fluoride and its associations with systemic health; it is important for the reader to be aware of these links when interpreting the aggregated evidence, so that the same data are not counted more than once.

4.5 Future research

High-quality prospective longitudinal studies to monitor the effect of individual-level fluoride exposures (including artificially fluoridated drinking water, other liquids, foods, toothpastes and mouthwashes, and fluoride-based dental interventions) on systemic health (including neuropsychological conditions, endocrine conditions, cancers, bone health and fractures, renal conditions, and birth abnormalities) are required In Ireland and other countries that implement CWF. Joint cross-country analyses should be conducted in order to increase the power of the findings. Such longitudinal studies must account for all potential confounding factors and effect modifiers (such as micronutrient deficiencies and excesses, where these are relevant).

Additionally, more accurate assessments of exposure to fluoride are required to better understand the dose-response relationship. Ecological assignment of fluoridation status based on region, or individual

assignment based on address history, only provides a crude indicator of exposure. Studies that measure fluoride concentrations in tap water samples and carefully assess the dose of exposure based on water and food consumption will provide more useful data. More precise data collection on sources of fluoride from seafood, tea, processed foods, fluoride supplements and toothpaste will also provide important context for the findings in relation to artificial CWF specifically.

Collaboration with public entities in Ireland and abroad may provide an important avenue for successful study implementation, particularly organisations involved in water treatment and distribution and large-scale cohort studies, into which investigations of fluoride exposure and possible effects could be embedded.

4.6 Conclusions

This review, encompassing 30 studies from nine countries, including Ireland, between 1990 and September 2021, indicates that there continues to be no definitive evidence that CWF has negative health effects. We found no conclusive evidence for a link between CWF and most conditions we examined for which research was available, including bone health, cancer, kidney stones, birth and infant abnormalities, and death rates. A summary of findings for each of the eight outcomes is displayed in Table 42.

While bone health and cancer have previously been primary areas of concern for researchers, the findings of this review point to generally mixed or null findings in relation to these outcomes. However, neuropsychological and endocrine outcomes emerged as areas requiring further monitoring; as the existing research in this area is currently limited in scope and interpretation is hampered by methodological problems, further high-quality research is now needed in order to shed light on the impact, if any, of artificial water fluoridation on these aspects of systemic health.

Table 42 Summary of conclusions of each outcome

Outcome	Conclusion	GRADE quality of evidence
Bone health	Mixed findings and many null findings around BMD and fractures; no association established, and evidence leans in favour of no association	Very low
Neuropsychological outcomes	Mixed findings on IQ and ADHD; methodological problems call validity of findings into question	Very low
Cancer	Evidence points to no association with artificially fluoridated water	Very low
Endocrine conditions	Mixed findings on associations with artificially fluoridated water from a small number of ecological and cross-sectional studies; future studies must account for iodine	Very low
Renal conditions	Findings from only one ecological study (protective)	Very low
Birth or birthing abnormalities	No association with artificially fluoridated water for a range of outcomes; ecological studies only	Very low
Infant abnormalities	Findings from only one case-control study (null)	Very low
All-cause mortality	Findings from only one ecological study (protective)	Very low

References

1 British Geological Survey. Fluoride in groundwater. BGS Br. Geol. Soc. 2021.www2.bgs.ac.uk/groundwater/health/fluoride.html (accessed 3 Sep 2021).

2 World Health Organization. *Guidelines for drinking-water quality: Fourth edition incorporating the first addendum*. 1st add, 4th ed. Geneva, Switzerland: World Health Organisation 2017. apps.who.int/iris/handle/10665/254636 (accessed 2 Sep 2021).

3 Public Health England. Water fluoridation: Health monitoring report for England 2014. London: Public Health England 2014. www.gov.uk/government/publications/water-fluoridation-health-monitoring-report-for-england-2014 (accessed 4 Aug 2021).

4 Arvin E, Bardow A, Spliid H. Caries affected by calcium and fluoride in drinking water and family income. *J Water Health* 2018;**16**:49–56. doi:10.2166/wh.2017.139

5 McDonagh MS, Whiting PF, Wilson PM, *et al.* Systematic review of water fluoridation. *BMJ* 2000;**321**:855. doi:10.1136/bmj.321.7265.855

6 Canadian Agency for Drugs and Technologies in Health. Community Water Fluoridation Programs: A Health Technology Assessment — Review of Dental Caries and Other Health Outcomes. Ottawa: Canadian Agency for Drugs and Technologies in Health 2019. www.cadth.ca/community-water-fluoridation-programs-health-technology-assessment (accessed 2 Sep 2021).

7 Iheozor-Ejiofor Z, Worthington HV, Walsh T, *et al.* Water fluoridation for the prevention of dental caries. *Cochrane Database Syst Rev* 2015;**2015**:263. doi:10.1002/14651858.CD010856.pub2

8 Singh KA, Spencer AJ, Armfield JM. Relative Effects of Pre- and Posteruption Water Fluoride on Caries Experience of Permanent First Molars. *J Public Health Dent* 2003;**63**:11–9. doi:10.1111/j.1752-7325.2003.tb03469.x

9 World Health Organization. WHO Guidelines for Drinking Water Quality Training Pack. Geneva, Switzerland: World Health Organization 2000. apps.who.int/iris/handle/10665/66218 (accessed 22 Nov 2021).

10 World Health Organization. Fluoride in Drinking-water: Background document for development of WHO Guidelines for Drinking-water Quality. World Health Organization 2004. www.who.int/water_sanitation_health/dwq/chemicals/fluoride.pdf (accessed 1 Sep 2021).

11 Peckham S, Lowery D, Spencer S. Are fluoride levels in drinking water associated with hypothyroidism prevalence in England? A large observational study of GP practice data and fluoride levels in drinking water. *J Epidemiol Community Health* 2015;**69**:619–24. doi:10.1136/jech-2014-204971

12 Harrison PT. Fluoride in water: A UK perspective. *J Fluor Chem* 2005;**126**:1448–56. doi:10.1016/j.jfluchem.2005.09.009

13 Peckham S, Awofeso N. Water Fluoridation: A Critical Review of the Physiological Effects of Ingested Fluoride as a Public Health Intervention. *Sci World J* 2014;**2014**:1–10. doi:10.1155/2014/293019

14 British Fluoridation Society. The extent of water fluoridation. Oldham, UK: British Fluoridation Society 2012. bfsweb.org/one-in-a-million/ (accessed 2 Sep 2021).

15 Johnston NR, Strobel SA. Principles of fluoride toxicity and the cellular response: a review. *Arch Toxicol* 2020;**94**:1051–69. doi:10.1007/s00204-020-02687-5

16 Government of Ireland. Health (Fluoridation of Water Supplies) Act, 1960. 1960. https://www.irishstatutebook.ie/eli/1960/act/46/enacted/en/html (accessed 2 Sep 2021).

17 Whelton H, Crowley E, Harding M, *et al.* North South survey of children's oral health in Ireland 2002. Brunswick Press Ltd 2002. www.lenus.ie/handle/10147/119028 (accessed 20 Oct 2021).

18 Government of Ireland. Forum on Fluoridation. Dublin: Government of Ireland 2002. www.lenus.ie/handle/10147/42506 (accessed 9 Feb 2021).

19 Government of Ireland. S.I. No. 42/2007 - Fluoridation of Water Supplies Regulations 2007. 2007. www.irishstatutebook.ie/2007/en/si/0042.html (accessed 2 Sep 2021).

20 Comber H, Deady S, Montgomery E, *et al.* Drinking water fluoridation and osteosarcoma incidence on the island of Ireland. *Cancer Causes Control* 2011;**22**:919–24. doi:10.1007/s10552-011-9765-0

21 Cronin J, Moore S, Harding M, *et al.* A cost-effectiveness analysis of community water fluoridation for schoolchildren. *BMC Oral Health* 2021;**21**:158. doi:10.1186/s12903-021-01490-7

22 National Health and Medical Research Council. A Systematic Review of the Efficacy and Safety of Fluoridation. Sydney, Canberra: Australian Government 2007. www.nhmrc.gov.au/sites/default/files/documents/attachments/A-systematic-review-of-the-Efficacy-and-safety-of-Fluoridation-part-a%20.pdf (accessed 4 Jul 2021).

23 University of Toronto. Water fluoridation: questions & answers. Toronto, ON: University of Toronto 2012. www.caphd.ca/sites/default/files/WaterFluoridationQA.pdf (accessed 2 Sep 2021).

24 Cheng K, Chalmers I, Sheldon TA. Adding fluoride to water supplies. *BMJ* 2007;**335**:699–702. doi:10.1136/bmj.39318.562951.BE

25 Sutton M, Kiersey R, Farragher L, *et al.* Health effects of water fluoridation. Dublin: Health Research Board 2015. www.hrb.ie/fileadmin/publications_files/Health_Effects_of_Water_Fluoridation.pdf (accessed 4 Jul 2021).

26 James P, Harding M, Beecher T, *et al.* Impact of Reducing Water Fluoride on Dental Caries and Fluorosis. *J Dent Res* 2021;**100**:507–14. doi:10.1177/0022034520978777

27 Centre for Reviews and Dissemination. *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. York, UK: University of York 2008. www.york.ac.uk/media/crd/Systematic_Reviews.pdf

28 Page MJ, McKenzie JE, Bossuyt PM, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71. doi:10.1136/bmj.n71

29 Bramer WM, Rethlefsen ML, Kleijnen J, *et al.* Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev* 2017;**6**:1–12. doi:10.1186/s13643-017-0644-y

30 Suarez-Almazor ME, Flowerdew G, Saunders L, *et al.* The fluoridation of drinking water and hip fracture hospitalization rates in two Canadian communities. *Am J Public Health* 1993;**83**:689–93. doi:10.2105/ajph.83.5.689

31 Lehmann R, Wapniarz M, Hofmann B, *et al.* Drinking water fluoridation: Bone mineral density and hip fracture incidence. *Bone* 1998;**22**:273–8. doi:10.1016/s8756-3282(97)00273-1

32 Kröger H, Alhava E, Honkanen R, *et al.* The effect of fluoridated drinking water on axial bone mineral density — a population-based study. *Bone Min* 1994;**27**:33–41. doi:10.1016/s0169-6009(08)80184-5

33 Higgins J, Thomas J, Chandler J, *et al.*, editors. *Cochrane handbook for systematic reviews of interventions version 6.2*. John Wiley & Sons 2019. www.training.cochrane.org/handbook.

34 Pluye P, El Sherif R, Gore G, *et al.* Facilitating the identification of qualitative, quantitative and mixed methods studies: Pilot evaluation study of a database filter for mixed studies reviews. Boston, MA: Mixed Methods International Research Association (MMIRA) conference 2014. 2.toolkit4mixedstudiesreviews.pbworks.com/w/file/fetch/82632454/MEDLINE%20Mixed%20methods%20or%20quantitative%20or%20qualitative.pdf

35 McGowan J, Sampson M, Salzwedel D, *et al.* CADTH Methods and Guidelines: PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Explanation and Elaboration. *CADTH Ott Can* 2016;:79.

36 EPPI-Centre Software. EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis. 2020.

37 Haddaway NR, Grainger, Gray CT. Citationchaser. 2021.estech.shinyapps.io/citationchaser/?dois=10.1038/s41559-020-01295-x,10.1371/journal.pone.0138237

38 Ibarluzea J, Gallastegi M, Santa-Marina L, *et al.* Prenatal exposure to fluoride and neuropsychological development in early childhood: 1-to 4 years old children. *Env Res* 2021;:112181. doi:10.1016/j.envres.2021.112181

39 Hennekens CH, Buring J. *Epidemiology in Medicine*. 1st Edition. Philadelphia, PA: Lippincott Williams & Wilkins 1987.

40 National Heart Lung and Blood Institute. Study Quality Assessment Tools. Sydney, Canberra: 2021. www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools (accessed 21 Apr 2021).

41 Young N, Newton J, Morris J, *et al.* Community water fluoridation and health outcomes in England: a cross-sectional study. *Community Dent Oral Epidemiol* 2015;**43**:550–9. doi:10.1111/cdoe.12180

42 Quigley JM, Woods B, Thompson J, *et al.* PRM209 - Confronting heterogeneity: Using systematic review effectively for meta-analysis. In: *Value Health*. 2013. A323–636. doi:doi: 10.1016/j.jval.2013.08.1763

43 Cope S, Zhang J, Saletan S, *et al.* A process for assessing the feasibility of a network meta-analysis: A case study of everolimus in combination with hormonal therapy versus chemotherapy for advanced breast cancer. *BMC Med* 2014;**12**:93. doi:10.1186/1741-7015-12-93

44 Schünemann H, Brożek J, Guyatt G, *et al. The GRADE Handbook*. Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group 2013. gdt.gradepro.org/app/handbook/handbook.html (accessed 4 Jul 2021).

45 Cauley J, Murphy P, Riley T, *et al.* Effects of fluoridated drinking water on bone mass and fractures: the study of osteoporotic fractures. *J Bone Min Res* 1995;**10**:1076–86. doi:10.1002/jbmr.5650100712

46 Phipps KR, Orwoll ES, Mason JD, *et al.* Community water fluoridation, bone mineral density, and fractures: prospective study of effects in older women. *BMJ* 2000;**321**:860–4. doi:10.1136/bmj.321.7265.860

47 Bassin EB, Wypij D, Davis RB, *et al.* Age-specific fluoride exposure in drinking water and osteosarcoma (United States). *Cancer Causes Control* 2006;**17**:421–8. doi:10.1007/s10552-005-0500-6

48 Kim FM, Hayes C, Burgard S, *et al.* A Case-Control Study of Fluoridation and Osteosarcoma. *J Dent Res* 2020;**99**:1157–64. doi:10.1177/0022034520919385

49 Barberio AM, Quiñonez C, Hosein FS, *et al.* Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation. *Can J Public Health* 2017;**108**:e229–39. doi:10.17269/cjph.108.5951

50 Riddell JK, Malin AJ, Flora DB, *et al.* Association of water fluoride and urinary fluoride concentrations with attention deficit hyperactivity disorder in Canadian youth. *Env Int* 2019;**133**:105190–105190. doi:10.1016/j.envint.2019.105190

51 Barberio AM, Hosein FS, Quiñonez C, *et al.* Fluoride exposure and indicators of thyroid functioning in the Canadian population: implications for community water fluoridation. *J Epidemiol Community Health* 2017;**71**:1019–25. doi:10.1136/jech-2017-209129

52 Cunningham JEA, McCague H, Malin AJ, *et al.* Fluoride exposure and duration and quality of sleep in a Canadian population-based sample. *Env Health* 2021;**20**:1–10. doi:10.1186/s12940-021-00700-7

53 Malin AJ, Riddell J, McCague H, *et al.* Fluoride exposure and thyroid function among adults living in Canada: Effect modification by iodine status. *Env Int* 2018;**121**:667–74. doi:10.1016/j.envint.2018.09.026

54 Green R, Lanphear BP, Hornung R, *et al.* Association Between Maternal Fluoride Exposure During Pregnancy and IQ Scores in Offspring in Canada. *JAMA Pediatr* 2019;**173**:940–8. doi:10.1001/jamapediatrics.2019.1729

55 Till C, Green R, Flora D, *et al.* Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. *Env Int* 2020;**134**:105315. doi:10.1016/j.envint.2019.105315

56 Farmus L, Till C, Green R, *et al.* Critical Windows of Fluoride Neurotoxicity in Canadian Children. *Env Res* 2021;**200**:111315–111315. doi:10.1016/j.envres.2021.111315

57 Arnold CM, Bailey DA, Faulkner RA, *et al.* The effect of water fluoridation on the bone mineral density of young women. *Can J Public Health* 1997;**88**:388–91. doi:10.1007/bf03403912

58 Blakey KF Richard G; Parslow, Roger; James, Peter W; Pozo, B G; Stiller, Charles A; Vincent, Tim J; Norman, Paul; McKinney, P A; Murphy, Michael F; Craft, Alan W; McNally, Richard JQ. Is fluoride a risk factor for bone cancer? Small area analysis of osteosarcoma and Ewing sarcoma diagnosed among 0–49-year-olds in Great Britain, 1980–2005. *Int J Epidemiol* 2014;**43**:224–34. doi:10.1093/ije/dyt259

59 Broadbent JM; T W Murray; Ramrakha, Sandhya; Moffitt, Terrie E; Zeng, Jiaxu; Page, Lyndie A Foster; Poulton, Richie. Community Water Fluoridation and Intelligence: Prospective Study in New Zealand. *Am J Public Health* 2015;**105**:72–6. doi:10.2105/ajph.2013.301857

60 Chachra D, Limeback H, Willett TL, *et al.* The long-term effects of water fluoridation on the human skeleton. *J Dent Res* 2010;**89**:1219–23. doi:10.1177/0022034510376070

61 Cohn P. An epidemiological report on drinking water. Trenton, NJ: New Jersey Department of Health, Environmental Health Service 1992.

62 Crnosija N, Choi M, Meliker JR. Fluoridation and county-level secondary bone cancer among cancer patients 18 years or older in New York State. *Env Geochem Health* 2019;**41**:761–8. doi:10.1007/s10653-018-0170-4

63 Danielson C, Lyon J, Egger M, *et al.* Hip fractures and fluoridation in Utah's elderly population. *JAMA* 1992;**268**:746–8. doi:10.1001/jama.1992.03490060078027

64 Dick A, Ford R, Schluter P, *et al.* Water fluoridation and the sudden infant death syndrome. *N Z Med J* 1999;**112**:286–9.

65 Hrudey SE; S Colin L; Berkel, J; Fincham, S. Drinking water fluoridation and osteosarcoma. *Can J Public Health* 1990;**81**:415–6.

66 Jacobsen SJ, Goldberg J, Cooper C, *et al.* The association between water fluoridation and hip fracture among white women and men aged 65 years and older. A national ecologic study. *Ann Epidemiol* 1992;**2**:617–26. doi:10.1016/1047-2797(92)90006-c

67 Jacobsen SJ, O'Fallon W, Melton LJ. Hip fracture incidence before and after the fluoridation of the public water supply, Rochester, Minnesota. *Am J Public Health* 1993;**83**:743–5. doi:10.2105/ajph.83.5.743

68 Lee N, Kang S, Lee W, *et al.* The Association between Community Water Fluoridation and Bone Diseases: A Natural Experiment in Cheongju, Korea. *Int J Env* 2020;**17**:9170-NA. doi:10.3390/ijerph17249170

69 Lowry R, Steen N, Rankin J. Water fluoridation, stillbirths, and congenital abnormalities. *J Epidemiol Community Health* 2003;**57**:499–500. doi:10.1136/jech.57.7.499

70 Mahoney M, Nasca P, Burnett W, *et al.* Bone cancer incidence rates in New York State: time trends and fluoridated drinking water. *Am J Public Health* 1991;**81**:475–9. doi:10.2105/ajph.81.4.475

71 McGuire SM, Vanable ED, McGuire MH, *et al.* Is There a Link Between Fluoridated Water and Osteosarcoma. *J Am Dent Assoc* 1991;**122**:38–45. doi:10.14219/jada.archive.1991.0149

72 National Fluoridation Information Service. Community Water Fluoridation Community Water Fluoridation and Osteosarcoma - Evidence from Cancer Registries Evidence from Cancer Registries. Lower Hutt, NZ: National Fluoridation Information Service 2013. www.moh.govt.nz/NoteBook/nbbooks.nsf/0/4C5B112CE74F7279CC257E49007A28EB/$file/Community%20Water%20Fluoridation%20and%20Osteosarcoma.pdf (accessed 4 Jul 2021).

73 O'Sullivan V, O'Connell BC. Water fluoridation, dentition status and bone health of older people in Ireland. *Community Dent Oral Epidemiol* 2014;**43**:58–67. doi:10.1111/cdoe.12130

74 Zhang X, Lu E, Stone SL, *et al.* Dental Cleaning, Community Water Fluoridation and Preterm Birth, Massachusetts: 2009-2016. *Matern Child Health J* 2019;**23**:451–8. doi:10.1007/s10995-018-2659-y

75 Vestergaard P, Jorgensen NR, Schwarz P, *et al.* Effects of treatment with fluoride on bone mineral density and fracture risk - a meta-analysis. *Osteoporos Int* 2008;**19**:257–68. doi:10.1007/s00198-007-0437-6

76 Søgaard CH, Mosekilde Li, Richards A, *et al.* Marked decrease in trabecular bone quality after five years of sodium fluoride therapy; assessed by biomechanical testing of iliac crest bone biopsies in osteoporotic patients. *Bone* 1994;**15**:393–9. doi:10.1016/8756-3282(94)90815-X

77 Allen MR, Krohn K. Skeletal Imaging. In: *Basic and Applied Bone Biology*. Elsevier 2014. 93–113. doi:10.1016/B978-0-12-416015-6.00004-6

78 Neer R. The utility of single-photon absorptiometry and dual-energy x-ray absorptiometry. *J Nucl Med* 1992;**33**:170–1.

79 Ilich JZ, Hsieh LC, Tzagournis MA, *et al.* A comparison of single photon and dual X-ray absorptiometry of the forearm in children and adults. *Bone* 1994;**15**:187–91. doi:10.1016/8756-3282(94)90706-4

80 Weinstein RS, New KD, Sappington LJ. Dual-energy X-ray absorptiometry versus single photon absorptiometry of the radius. *Calcif Tissue Int* 1991;**49**:313–6. doi:10.1007/BF02556252

81 Nelson D, Feingold M, Mascha E, *et al.* Comparison of single-photon and dual-energy x-ray absorptiometry of the radius. *Bone Miner* 1992;**18**:77–83. doi:10.1016/0169-6009(92)90801-J

82 Guth S, Hüser S, Roth A, *et al.* Toxicity of fluoride: critical evaluation of evidence for human developmental neurotoxicity in epidemiological studies, animal experiments and in vitro analyses. *Arch Toxicol* 2020;**94**:1375–415. doi:10.1007/s00204-020-02725-2

83 Miranda GHN, Alvarenga MOP, Ferreira MKM, *et al.* A systematic review and meta-analysis of the association between fluoride exposure and neurological disorders. *Sci Rep* 2021;**11**:22659. doi:10.1038/s41598-021-99688-w

84 National Toxicology Program. Draft NTP Monograph on the Systematic Review of Fluoride Eposure and Neurodevelopmental and Cognitive Health Effects. National Toxicology Program, US Department of Health and Human Services 2019.

85 González G, Karipidis I, Tijms J. Dyslexia as a Neurodevelopmental Disorder and What Makes It Different from a Chess Disorder. *Brain Sci* 2018;**8**:189. doi:10.3390/brainsci8100189

86 Freire C, Amaya E, Gill F, *et al.* Prenatal co-exposure to neurotoxic metals and neurodevelopment in preschool children: The Environment and Childhood (INMA) Project. *Sci Total Environ* 2018;**621**:340–51.

87 Yiamouyiannis J, Burk D. Fluoridation and cancer age-dependence of cancer mortality related to artificial fluoridation. *Natl Cancer Program Fluorid Public Drink Water* 1977;:19.

88 National Research Council. Health Effects of Ingested Fluoride. Washington, DC: National Research Council 1993.

89 Wu L, Fan C, Zhang Z, *et al.* Association between fluoride exposure and kidney function in adults: A cross-sectional study based on endemic fluorosis area in China. *Ecotoxicol Env Saf* 2021;**225**:112735. doi:10.1016/j.ecoenv.2021.112735

90 Agueda AR Josep Ma; Manau, Carolina; Guerrero, Adrian; Echeverría, José J. Periodontal disease as a risk factor for adverse pregnancy outcomes: a prospective cohort study. 2007;**0**:71127132039003-??? doi:10.1111/j.1600-051x.2007.01166.x

91 Heimonen A, Rintamäki H, Furuholm J, *et al.* Postpartum oral health parameters in women with preterm birth. *Acta Odontol Scand* 2008;**66**:334–41. doi:10.1080/00016350802307620

92 Fluegge K. Community water fluoridation predicts increase in age-adjusted incidence and prevalence of diabetes in 22 states from 2005 and 2010. *J Water Health* 2016;**14**:864–77. doi:10.2166/wh.2016.012

93 Schwartz GG. Eye Cancer Incidence in U.S. States and Access to Fluoridated Water. *Cancer Epidemiol Biomark Prev* 2014;**23**:1707–11. doi:10.1158/1055-9965.epi-14-0437

94 The Royal Society of New Zealand. Health effects of water fluoridation: A review of the scientific evidence. Auckland: The Royal Society of New Zealand, Office of the Prime Minister's Chief Science Advisor 2014. researcharchive.lincoln.ac.nz/bitstream/handle/10182/13121/Health-effects-of-water-fluoridation_Aug_2014.pdf?sequence=1 (accessed 30 Nov 2021).

95 Office of the Prime Minister's Chief Science Advisor. Fluoridation: An Update on Evidence. Auckland: Office of the Prime Minister's Chief Science Advisor, University of Auckland 2021. www.pmcsa.ac.nz/topics/fluoridation-an-update-on-evidence/ (accessed 7 Apr 2021).

96 National Health and Medical Research Council. Evaluating the Evidence on Water Fluoridation and Human Health in Australia 2014-2017: Administrative Report. Sydney, Canberra: Australian Government 2017. www.nhmrc.gov.au/about-us/publications/water-fluoridation-dental-and-other-human-health-outcomes (accessed 21 Apr 2021).

97 Douglass CW, Joshipura K. Caution needed in fluoride and osteosarcoma study. *Cancer Causes Control* 2006;**17**:481–2. doi:10.1007/s10552-006-0008-8

98 Kim FM, Hayes C, Williams PL, *et al.* An Assessment of Bone Fluoride and Osteosarcoma. *J Dent Res* 2011;**90**:1171–6. doi:10.1177/0022034511418828

Appendix A Epidemiology of outcomes

Bone health

Bone mineral density and osteoporosis

Definition: Osteoporosis, meaning “porous bone”, is a bone disease that occurs when the body loses too much bone, makes too little bone, or both. As a result, the bones becomes fragile and may break more easily. Research over more than 30 years has constistenty shown that both high dose and slow-release sodium fluoride increases spinal bone mass and may decrease the frequency of vertebral fractures.

Affected population: Osteoporosis is the major cause of fractures in the spine, hip, wrist, humerus and pelvis among postmenopausal women and in older men.

Predisposing or influencing factors: Factors that may increase the risk for osteoporosis include: female sex; age 70 or over; thin-boned women and men; white and Asian women (white men are at higher risk than African American and Mexican American men); family history of osteoporosis or hip fracture; low levels of certain hormones, such as low oestrogen levels in women after menopause or low levels of testosterone in men; a diet low in calcium and vitamin D; some medical conditions such as endocrine and hormonal diseases, gastrointestinal diseases, rheumatoid arthritis, certain types of cancer, HIV/AIDS, and anorexia nervosa; medications such as glucocorticoids and adrenocorticotropic hormone used to treat asrtritis and asthma, antiepileptic medicines, which treat seizures and other neurological disorders, cancer medications, which use hormones to treat breast and prostate cancer, proton pump inhibitors, which lower stomach acid, thiazolidinediones for the treatment of type II diabetes, and selective serotonin reuptake inhibitors, which treat depression and anxiety; lifestyle factors such as low levels of physical activity and lengthy periods of inactivity, alcoholism, and smoking. Researchers are still trying to asertain whether the influence of smoking on bone health is solely due to tobacco usage or whether smokers have other risk factors for osteoporosis.

Fluoride theory and evidence: Morphological examinations of bone biopsies have shown that fluoride treatment for osteoporosis primarily stimulates osteoblasts in cancellous bones. Although studies have repeatedly shown that fluoride treatment enhances spinal bone mass, no consistent decrease in the risk of vertebral fractures has been reported. It has been claimed that fluoride therapy may increase the risk of nonvertebral fractures, such as hip fractures. The use of sodium fluoride in the treatment of osteoporosis was a matter of debate in the 1990s.

Sources:

1. O’ Sullivan V, O’ Connell BC. Water fluoridation, dentition status and bone health of older people in Ireland. Community dentistry and oral epidemiology. 2015 Feb;43(1):58-67.
2. Lehmann R, Wapniarz M, Hofmann B, Pieper B, Haubitz I, Allolio B. Drinking water fluoridation: Bone mineral density and hip fracture incidence. Bone. 1998 Mar 1;22(3):273-8.
3. Osteoporosis Overview | NIH Osteoporosis and Related Bone Diseases National Resource Center [Internet]. Bones.nih.gov. 2021 [cited 13 July 2021]. Available from: https://www.bones.nih.gov/health-info/bone/osteoporosis/overview

Fracture

Definition: Fractures are cracks or breaks in bones, commonly caused by car accidents, falls, or sports injuries. They can also be caused by a health condition that weakens the bone, such as osteoporosis or cancer. Hip fractures, specifically, are cracks or breaks at the hip joint in the top of the thigh bone (femur).

Affected population: White women, followed by white males, are at the highest risk of fracture. The majority of these fractures occur as a result of osteoporosis in the elderly. However, researchers have expressed concern that one of the medical treatments for osteoporosis, oral fluoride consumption (75 mg/d), may increase the risk of hip fracture specifically. Public Health England identified the following confounding factors a priori: age (proportion of population above 65 years old); gender (proportion of the population female); high deprivation; ethnicity (proportion of the population white).

Predisposing or influencing factors: The major risk factors are advanced age, female gender, osteoporosis, smoking and alcohol consumption.

Fluoride theory and evidence: Among the environmental factors of bone metabolism postulated to account for geographic variation in hip fracture incidence, water fluoride content has received much attention. Water fluoridation can increase normal dietary intake of fluoride by around 50%, and about half of the fluoride consumed is absorbed by bone, which may have an effect on the mechanical characteristics of bones. Fluoridation having an influence on fracture risk is theoretically feasible.Fluoride stimulates the function of osteoblast-like cells in vitro. In humans, sodium fluoride administration enhances bone mineral density in the axial skeleton. Although most ecological studies failed to identify a preventive benefit for fractures, it was hypothesized that this added fluoride would strengthen bone. Sodium fluoride was first thought to be an effective therapy for osteoporosis. However, more recent research has found that sodium fluoride therapy increases the incidence of hip fracture. The increase in bone mass generated by fluoride has been linked to an increase in bone fragility, according to certain theories.

Sources:

1. Jacobsen SJ, Goldberg J, Cooper C, Lockwood SA. The association between water fluoridation and hip fracture among white women and men aged 65 years and older: a national ecologic study. Annals of epidemiology. 1992 Sep 1;2(5):617-26.
2. Jacobsen SJ, O'Fallon WM, Melton 3rd LJ. Hip fracture incidence before and after the fluoridation of the public water supply, Rochester, Minnesota. American journal of public health. 1993 May;83(5):743-5.
3. Danielson C, Lyon JL, Egger M, Goodenough GK. Hip fractures and fluoridation in Utah's elderly population. Jama. 1992 Aug 12;268(6):746-8.
4. Suarez-Almazor ME, Flowerdew G, Saunders LD, Soskolne CL, Russell AS. The fluoridation of drinking water and hip fracture hospitalization rates in two Canadian communities. American journal of public health. 1993 May;83(5):689-93.
5. Public Health England. Water fluoridation: Health monitoring report for England 2014. London: Public Health England; 2014.

Neuropsychological outcomes

IQ/neuropsychological development

Definition: An intelligence quotient (IQ) is a total score derived from a set of standardized tests or subtests designed to assess human intelligence. For modern IQ tests, the raw score is transformed to a normal distribution with mean 100 and standard deviation 15. As a result, approximately two-thirds of the population score between IQ 85 and IQ 115 while approximately 2.5% score above 130 and another 2.5% below 70. Since the early twentieth century, raw IQ scores for many populations have been increasing at an average rate of three IQ points each decade, a phenomenon known as the Flynn effect.

Predisposing or influencing factors: IQ scores have been linked to a variety of factors, including diet, lead or arsenic exposure, parental socioeconomic status, educational environment and opportunities, prenatal environment, and low birth weight. While the heritability of IQ has been studied for almost a century, disagreements persist concerning the importance of heritability estimates and inheritance processes. Variations in IQ performance across cultures and ethnic groups are often ascribed to cultural or linguistic bias rather than genetic differences.

Fluoride theory and evidence: The European Commission's Scientific Committee on Health and Environmental Risks reviewed studies published up to 2010 that examined the relationship between fluoride and intelligence and concluded that they used a simplistic methodological design with little control for confounding variables such as nutrition, iodine or lead exposure, or socioeconomic status. Fluoride crosses the placenta, and laboratory studies indicate that it accumulates in areas of the brain associated with learning and memory and affects proteins and neurotransmitters in the central nervous system. Higher fluoride exposure from drinking water has been linked to lower intelligience in a meta-analysis of 27 epidemiologic studies and investigations using biomarkers of fluoride exposure. All previous investigations, however, were cross-sectional in nature and were conducted in regions with higher water fluoride concentrations (0.88–31.6 mg/L) than those deemed acceptable in North America (0.7mg/L). Additionally, most studies did not assess exposure during foetal brain development. In a longitudinal birth cohort study involving 299 mother-child pairs in Mexico City (a region with naturally occurring fluoride and community salt-based fluoride), a 1-mg/L increase in maternal urinary fluoride (MUF) concentration was associated with a 6-point (95%CI, −10.84 to −1.74) lower IQ score among school-aged children.

Sources:

1. Broadbent JM, Thomson WM, Ramrakha S, Moffitt TE, Zeng J, Foster Page LA, Poulton R. Community water fluoridation and intelligence: prospective study in New Zealand. American journal of public health. 2015 Jan;105(1):72-6.
2. Green R, Lanphear B, Hornung R, Flora D, Martinez-Mier EA, Neufeld R, Ayotte P, Muckle G, Till C. Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada. JAMA pediatrics. 2019 Oct 1;173(10):940-8.
3. Christie D. Introduction to IQ testing. Psychiatry. 2005 Jun 1;4(6):22-5.
4. Braaten EB, Norman D. Intelligence (IQ) testing. Pediatrics in review. 2006 Nov 1;27(11):403.
5. Farmus L, Till C, Green R, Hornung R, Martinez-Mier EA, Ayotte P, Muckle G, Lanphear B, Flora D. Critical windows of fluoride neurotoxicity in Canadian children. Environmental Research. 2021 May 27:111315.

6. Till C, Green R, Flora D, Hornung R, Martinez-Mier EA, Blazer M, Farmus L, Ayotte P, Muckle G, Lanphear B. Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. Environment international. 2020 Jan 1;134:105315.

Attention deficit hyperactivity disorder (ADHD)

Definition: Attention deficit hyperactivity disorder (ADHD) is the most prevalent neurodevelopmental condition in children. It is characterised by persistent and maladaptive symptoms of inattention, impulsivity/hyperactivity or both that begin in infancy and may last into adulthood. These symptoms may have a significant impact on learning and academic progress, as well as social functioning. Children with ADHD may have trouble paying attention, forgetfulness, disorganisation, controlling impulsive behaviours (they may act without considering the consequences), or be overly active. ADHD is estimated to affect 7–8% of school-aged children and 4–5% of adults. ADHD is classified into three subtypes: inattentive type, hyperactive/impulsive type, and mixed type. Attention Deficit Disorder, no hyperactivity [ADD] is an older term for what is now described as Attention Deficit Hyperactivity Disorder [ADHD], inattentive type. This subtype of ADHD is characterized by disorganization, forgetfulness, and difficulty paying attention, but not by hyperactive or impulsive behaviour.

Affected population: Children and adults

Predisposing or influencing factors: Although the cause(s) and risk factors for ADHD are unclear, current research on twins' link genes for ADHD indicates that genetics play a significant role. In addition to genetics, scientists are studying other possible causes and risk factors including brain injury; exposure to environmental elements (e.g., lead) during pregnancy or at a young age; alcohol and tobacco use during pregnancy; preterm birth; and low birth weight. Research does not support the popularly held views that ADHD is caused by excessive sugar consumption, excessive television viewing, poor parenting, or social and environmental factors such as poverty or family dysfunction.

ADHD is considered to develop as a result of a combination of genetic and environmental factors, with numerous developmental neurotoxicants significantly increasing the risk for a diagnosis of ADHD. Environmental factors include prenatal and neonatal exposure to manganese, poly-chlorinated biphenyls, nicotine, and mercury, as well as childhood exposure to arsenic, food additives and food colouring, pesticides, and lead.

Huber *et al.* investigated a possible association of ADHD prevalence with altitude, hypothesising that mild hypobaric hypoxia at higher altitudes may increase dopamine levels. Based on findings of a number of studies that reported decreased dopamine levels in children and adolescents with ADHD, they hypothesised that higher-altitude areas in the United States may have a lower prevalence of ADHD. They considered a range of potential social or psychological risk-modifying factors in addition to altitude, but did not include CWF or potential chemical toxicants. Silicofluoride-treated water could corrode lead pipes laid in the 1930's; however, few countries have significant amounts of lead piping remaining in their water system.

Fluoride theory and evidence: Fluoride has received little attention in the ADHD literature, despite being environmentally widespread and having demonstrable developmental neurotoxic effects at a sufficient dose. The Perrott *et al.* (2015) critique briefly repeats the statistical analyses of Malin and Till (2015) and Huber *et al.* using the same data and data for other possible risk-modifying factors available from similar sources. The only covariates showing a statistically significant association were mean state elevation and per capita personal income in 2009. There are also purported links between ADHD and thyroid hormone disruption, which may mediate any effect of fluoride exposure on ADHD.

Sources:

1. Malin AJ, Till C. Exposure to fluoridated water and attention deficit hyperactivity disorder prevalence among children and adolescents in the United States: an ecological association. Environmental Health. 2015 Dec;14(1):1-0.
2. Perrott KW. Fluoridation and attention deficit hyperactivity disorder–a critique of Malin and Till (2015). British dental journal. 2017 Dec;223(11):819-22.
3. Antshel KM, Hargrave TM, Simonescu M, Kaul P, Hendricks K, Faraone SV. Advances in understanding and treating ADHD. BMC medicine. 2011 Dec;9(1):1-2.
4. What is ADHD? [Internet]. Centers for Disease Control and Prevention. Available from: https://www.cdc.gov/ncbddd/adhd/facts.html
5. Huber R S, Kim T-S, Kim N, *et al.* Association Between Altitude and Regional Variation of ADHD in Youth. J Atten Disord 2015; DOI: 10.1177/1087054715577137.

Learning disability diagnosis (Dyslexia)

Definition: Dyslexia is a specific learning disorder and neurodevelopmental disorder characterised by slow and inaccurate word recognition. Individuals with dyslexia have difficulties accurately and fluently recognising and spelling words, despite adequate instruction and intelligence and intact sensory abilities. Prevalence is approximately 7% of the population, according to one common definition (reading achievement 1.5 standard deviations below the mean for age).

Affected population: Children and adults

Predisposing or influencing factors: Dyslexia is a genetic disorder that is moderately heritable. It is also associated with multiple environmental risk factors, including maternal smoking during the first three years of the child's life, the child's literacy-related activity, and use of electronic devices.

Fluoride theory and evidence: Histological, chemical, and molecular research have demonstrated a biologically plausible link between fluoride and altered brain function. Second, clarifying the nature of this relationship is important and timely, because fluoride was recently classified as one of six new neurodevelopmental toxins.

Sources:

1. Barberio AM, Quinonez C, Hosein FS, McLaren L. Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation. Canadian Journal of Public Health. 2017 May;108(3):e229-39.
2. Antshel KM, Hargrave TM, Simonescu M, Kaul P, Hendricks K, Faraone SV. Advances in understanding and treating ADHD. BMC medicine. 2011 Dec;9(1):1-2.
3. Becker N, Vasconcelos M, Oliveira V, Santos FC, Bizarro L, Almeida RM, Salles JF, Carvalho MR. Genetic and environmental risk factors for developmental dyslexia in children: Systematic review of the last decade. Developmental neuropsychology. 2017 Dec 1;42(7-8):423-45.
4. Peterson RL, Pennington BF. Developmental dyslexia. The lancet. 2012 May 26;379(9830):1997-2007.

Cancer

Osteosarcoma

Definition: Osteosarcoma is a very rare primary malignant tumour of bone. Approximately, 80% of these tumours tend to occur in the long bones of the appendicular skeleton usually when undergoing periods of rapid growth. The disease is caused by a malignant transformation of bone-forming cells and is most prevalent in late adolescence. Osteosarcoma begins when a healthy bone cell develops changes in its DNA.

Affected population: Two age cohorts are affected by osteosarcoma and the incidence of osteosarcoma is slightly higher in males than females. The young cohort are mostly male and aged 15–20 years old (or under 25 years in other articles). The second cohort is composed mostly of men aged 60 and above. Paget's disease is associated with osteosarcoma is the older age cohort. Paget's disease prevalence and osteosarcoma incidence in elderly patients have similar geographical distributions. Paget's disease of the bone interferes with the body's normal recycling process, in which new bone tissue gradually replaces old bone tissue. Over time, bones can become fragile and misshapen. The pelvis, skull, spine and legs are most commonly affected.

Predisposing or influencing factors: The factors that increase the risk of osteosarcoma include previous radiation therapy; other bone disorders (such as Paget's disease of bone and fibrous dysplasia, and certain inherited or genetic conditions, such as hereditary retinoblastoma, Bloom syndrome, Li-Fraumeni syndrome, Rothmund-Thomson syndrome, and Werner syndrome); ingestion of radioisotopes, high dose X-rays. Other risk factors include viral diseases, chemicals (Beryllium, vinyl chloride, and fluoride) and genetic susceptibility (rare genetic disorders). In humans, ionizing radiation is the only environmental agent known to cause bone cancer and is thought to effect approximately 3% of cases from either external high-dose irradiation used in cancer therapy or internal bone-seeking radioisotopes from occupational or medical use. Alkylating agents used in chemotherapy are thought to increase the risk for osteosarcoma and evidence for other etiologic factors including viruses, antecedent trauma, or radium in drinking water has been suggested but are inconclusive. Certain pre-existing bone defects such as Paget's disease, have been shown to be more prevalent in patients who subsequently developed bone cancers. Also, a genetic predisposition for osteosarcoma has been described, specifically for patients with a hereditary form of retinoblastoma or those with familial Li-Fraumeni cancer syndrome. A small number of research has also found renal metastasis among people with osteosarcoma.

Fluoride theory and evidence: Theoretical plausibility arises from deposition of fluoride in bone and a mitogenic effect on osteoblasts. There are no published findings on the molecular basis for the beginning of an osteosarcoma tumor in response to fluoride exposure in a refereed journal. However, the mechanism of fluoride's action to strengthen osteal tissue is related to its ionic capabilities. The major component of osteal tissue and the major strengthening material in teeth is structural hydroxyapatite, Ca5(P04)30H. When ionic fluoride enters the hydroxyapatite lattice, a dynamic exchange of fluoride for hydroxide occurs in portions of the structure. The resulting fluorapatite tightly binds the fluoride, strengthening the structure and rendering it less susceptible to dissolution in organic acids. Strengthening may occur by incorporation of fluoride ions during the tissue mineralization phase of tooth or bone growth, or via ion exchange after formation has occured.

Given that fluoride may act as a mitogen (increasing osteoblast proliferation) and that its uptake in bone increases with increased skeletal growth, it is biologically plausible that exposure to fluoride during growth may be associated with the subsequent development of osteosarcoma, and that fluoride may either increase or decrease the rate of osteosarcoma.

Sources:

1. Public Health England. Water fluoridation: health monitoring report for England 2014.
2. Hrudey SE, Soskolne CL, Berkel JO, Fincham SH. Drinking water fluoridation and osteosarcoma. Canadian Journal of Public Health. 1990;81(6):415-6.
3. McGuire SM, Vanable ED, McGuire MH, Buckwalter JA, Douglass CW. Is there a link between fluoridated water and osteosarcoma?. The Journal of the American Dental Association. 1991 Apr 1;122(4):38-45.
4. Cohn PD. An epidemiologic report on drinking water and fluoridation. New Jersey, Department of Health, Environmental Health Service; 1992.
5. Bassin EB, Wypij D, Davis RB, Mittleman MA. Age-specific fluoride exposure in drinking water and osteosarcoma (United States). Cancer Causes & Control. 2006 May;17(4):421-8.
6. Comber H, Deady S, Montgomery E, Gavin A. Drinking water fluoridation and osteosarcoma incidence on the island of Ireland. Cancer Causes & Control. 2011 Jun;22(6):919-24.
7. Blakey K, Feltbower RG, Parslow RC, James PW, Gómez Pozo B, Stiller C, Vincent TJ, Norman P, McKinney PA, Murphy MF, Craft AW. Is fluoride a risk factor for bone cancer? Small area analysis of osteosarcoma and Ewing sarcoma diagnosed among 0–49-year-olds in Great Britain, 1980–2005. International Journal of Epidemiology. 2014 Feb 1;43(1):224-34.
8. National Fluoridation Information Service (NFIS), 2013. Community Water Fluoridation Community Water Fluoridation and Osteosarcoma - Evidence from Cancer Registries. Available at: http://www.rph.org.nz/content/45a27238-502a-48ed-a90c-aa306b3ac449.cmr.
9. Akasbi Y, Arifi S, Lahlaidi K, Namad T, Mellas N, El Fassi MJ, Farih MH, Amarti A, El Mesbahi O. Renal metastases of a femur osteosarcoma: a case report and a review of the literature. Case reports in urology. 2012 Feb 28;2012.

Ewing sarcoma

Definition: Ewing sarcoma is a rare type of cancer that affects bones or surrounding tissue. It mainly affects children and young people but is also seen in adults, and more commonly in men than women. The DNA alterations that occur most often in Ewing sarcoma impact a gene called EWSR1.

Affected population: Ewing sarcoma is most likely to occur among children aged 10–24 years, men, and those of European ancestry.

Predisposing or influencing factors: Although the cause is unknown, genetic, and environmental factors are likely to be involved. Ewing sarcoma was associated with low population density and high levels of impoverishment in certain geographic areas.

Fluoride theory and evidence: No information provided to explain a theoretical association

Sources:

1. Blakey K, Feltbower RG, Parslow RC, James PW, Gómez Pozo B, Stiller C, Vincent TJ, Norman P, McKinney PA, Murphy MF, Craft AW. Is fluoride a risk factor for bone cancer? Small area analysis of osteosarcoma and Ewing sarcoma diagnosed among 0–49-year-olds in Great Britain, 1980–2005. International Journal of Epidemiology. 2014 Feb 1;43(1):224-34.

Bladder cancer

Definition: Bladder cancer occurs as a result of abnormal growth of cells in the lining of the bladder. There are three main types of bladder cancer

- Urothelial carcinoma (previously called transitional cell carcinoma), is a kind of bladder cancer that develops in the urothelial cells found in the urinary tract. Urothelial cells expand when the bladder is full and contract when the bladder is empty. Urothelial carcinoma is the most common type of bladder cancer in the USA.
- Squamous cell bladder cancer is associated with chronic irritation of the bladder, which may occur as a result of infection or long-term use of a urinary catheter. Squamous cell bladder cancer is rare in the USA but is more common in parts of the world where a certain parasitic infection (schistosomiasis) is a common cause of bladder infections.
- Adenocarcinoma begins in cells that make up mucus-secreting glands in the bladder. Adenocarcinoma of the bladder is very rare.

Affected population: Males, older age (>55 years), smoking, occupational exposure to aromatic amines and polycyclic aromatic hydrocarbons, genetic predisposition, received chemotherapy or radiation therapy for cancer, long-term catheter users, and those living in endemic schistosomiasis areas.

Predisposing or influencing factors: Factors that may increase bladder cancer risk include smoking; increasing age, as most people diagnosed with bladder cancer are older than 55; being male; exposure to certain chemicals including arsenic and chemicals used in the manufacture of dyes, rubber, leather, textiles, and paint products; previous cancer treatment, (e.g. cyclophosphamide or radiation treatments aimed at the pelvis); chronic bladder inflammation; and personal or family history of cancer. A family history of Lynch syndrome, also known as hereditary nonpolyposis colorectal cancer, can increase the risk of cancer in the urinary system, as well as in the colon, uterus, ovaries, and other organs.

Fluoride theory and evidence: Theoretical plausibility arises because fluoride is excreted in the urine and the bladder lining is therefore exposed to relatively high fluoride concentrations.

Sources:

1. Burger M, Catto JW, Dalbagni G, Grossman HB, Herr H, Karakiewicz P, Kassouf W, Kiemeney LA, La Vecchia C, Shariat S, Lotan Y. Epidemiology and risk factors of urothelial bladder cancer. European urology. 2013 Feb 1;63(2):234-41.

Secondary bone cancer

Definition: Secondary bone cancer is also known as bone metastasis. It occurs when a cancer that started somewhere else in the body spreads to the bones.

Affected population: Almost any type of cancer can spread to the bone. However, it is most likely to be caused by breast, kidney, lung or thyroid cancer.

Predisposing or influencing factors: No information available.

Fluoride theory and evidence: Given that secondary bone cancer is a relatively common metastatic site, it is theoretically possible that fluoride might influence bone's structural fortitude and therefore play a role in secondary cancer's spread to the skeleton.

Sources:

1. Crnosija N, Choi M, Meliker JR. Fluoridation and county-level secondary bone cancer among cancer patients 18 years or older in New York State. Environmental geochemistry and health. 2019 Apr;41(2):761-8.

All cancers

Definition: Cancer refers to a group of diseases characterised by abnormal cell growth, which may occur in almost any tissue in the body. Cancer cells may migrate from the site of origin, invading nearby tissues and forming masses of malignant cells, or tumours, at distant secondary sites in the body; this spreading is called metastasis. These tumours may become more aggressive over time and disrupt tissues and organs in the body, becoming potentially fatal. The most common sites for cancer are the breast, lungs, colon and rectum, prostate, skin, and stomach.

Affected population: The most common cancers occur most often in older people; the median age at time of a cancer diagnosis is 66.

Predisposing or influencing factors: Tobacco and alcohol use, obesity, low fruit and vegetable intake, lack of physical activity, certain chronic infections (e.g. hepatitis, human papillomavirus, Epstein-Barr virus)**.**

Fluoride theory and evidence: The Medical Research Council report (2002) concluded that the evidence available has not established that fluoride is genotoxic to humans and most of the studies suggest that it is not, but the possibility of some genotoxic effect cannot be excluded

Sources:

1. Public Health England. Water fluoridation: health monitoring report for England 2014.
2. Weinberg RA. How cancer arises. Scientific American. 1996 Sep 1;275(3):62-70.
3. Cancer. [Internet] World Health Organization [Updated 2021 September 21; cited 2021 November 10]. Available from: https://www.who.int/en/news-room/fact-sheets/detail/cancer

Endocrine conditions

Hypothyroidism and levels of thyroid-stimulating hormones (TSH)

Definition: Thyroid underactivity (hypothyroidism) is most commonly caused by an autoimmune disease known as Hashimoto's thyroiditis but can also occur due to the use of certain medications (e.g., lithium) and both increased and decreased iodine intake. Hypothyroidism, the most common thyroid disorder, is characterised by suppression of thyroid gland activity. Subclinical hypothyroidism is indicated by high serum thyroid stimulating hormone (TSH) concentrations of 4.5–9 mIU/L with normal triiodothyronine (T3) and thyroxine (T4) levels. However, TSH levels above 2.5 mIU/L may increase risk for subclinical and clinical hypothyroidism. Subclinical hypothyroidism is estimated to occur in 4.3–9.5% of the USA adult population and is associated with various health conditions. The prevalence of overt hypothyroidism in the general population varies between 0·3% and 3·7% in the USA and between 0·2% and 5·3% in Europe, depending on the definition used. A meta-analysis of studies across nine European countries estimated the prevalence of undiagnosed hypothyroidism, including both overt and mild cases, at around 5%.

Affected population: Women, age over 65 years, family history, white ancestry, people with iodine deficiency, people with other autoimmune conditions (type 1 diabetes, coeliac disease), people received radiation exposure or therapy. Differences in iodine status affect the prevalence of hypothyroidism, which

occurs more frequently both in populations with a relatively high iodine intake and in severely iodine-deficient populations.

Predisposing or influencing factors: Risk factors for thyroid diseases include, but are not limited to sex (female), age (over 50 or 60), family history of thyroid disease, and radiation exposure to the head or neck. Additionally, individuals with one autoimmune condition are more susceptible to developing other autoimmune conditions. For example, individuals with coeliac disease or diabetes have been found to have significantly higher risk of developing autoimmune thyroid diseases. Other risk factors are recent thyroid surgery (partial thyroidectomy) and been pregnant or delivered a baby within the past six months. Iodine deficiency is another causal factor. People with Down syndrome or Turners' syndrome have a higher risk of hypothyroidism. Certain drugs can induce hypothyroidism —for example, amiodarone, lithium, tyrosine kinase inhibitors, interferon-alfa, thalidomide, monoclonal antibody therapy (ipilimumab and nivolumab), antiepileptic drugs (valproate), and drugs for second-line treatment of multidrug-resistant tuberculosis. Fluoride exposures of 0.05–0.13 mg/kg/day have been associated with adverse thyroid effects among iodine sufficient people, while lower fluoride exposures of 0.01–0.03 mg/kg/day have been associated with these effects among iodine deficient people. Selenium deficiency may also be related to hypothyroidism, and this is not fully investigated.

Fluoride theory and evidence: The effects of fluoride on the thyroid have long been observed. In the 1950s, fluoride was used pharmacologically to suppress thyroid activity in people with hyperthyroidism. Doctors chose fluoride as a thyroid suppressor based on study findings linking fluoride to goitre, and fluoride therapy did reduce thyroid activity in the treated patients. Typically, a dose of between 2 and 5 mg fluoride per day was found to be effective and this is within the range commonly consumed by individuals living in fluoridated areas. Two reviews have examined the impact of fluoride on thyroid function, concluding that fluoride is an endocrine disruptor with the potential to disrupt the function of tissues that require iodine. It was suggested that the chief endocrine effect is decreased thyroid function at fluoride exposure levels as low as 0.01 mg/kg/day where iodine intake is inadequate. Iodine deficiency can contribute to decreased thyroid hormone production and exacerbate the thyroid-disrupting effects of certain chemicals, as well as fluoride. Adequate iodine levels can offset adverse goitrogenic effects of fluoride. Selenium deficiency may also be related to hypothyroidism, although this has not been fully investigated.

Sources:

1. Peckham S, Lowery D, Spencer S. Are fluoride levels in drinking water associated with hypothyroidism prevalence in England? A large observational study of GP practice data and fluoride levels in drinking water. J Epidemiol Community Health. 2015 Jul 1;69(7):619-24
2. Barberio AM, Hosein FS, Quiñonez C, McLaren L. Fluoride exposure and indicators of thyroid functioning in the Canadian population: implications for community water fluoridation. J Epidemiol Community Health. 2017 Oct 1;71(10):1019-25.
3. Malin AJ, Riddell J, McCague H, Till C. Fluoride exposure and thyroid function among adults living in Canada: effect modification by iodine status. Environment international. 2018 Dec 1;121:667-74.
4. Pearce EN. Is fluoridated drinking water associated with increased hypothyroidism risk?. Clinical Thyroidology. 2015 Apr 1;27(4):100-1.
5. Chaker L, Bianco AC, Jonklaas J, Peeters RP. Hypothyroidism. Lancet. 2017 Sep 23;390(10101):1550-1562. doi: 10.1016/S0140-6736(17)30703-1. Epub 2017 Mar 20. PMID: 28336049; PMCID: PMC6619426.

Sleep disturbance due to reduced functioning of pineal gland

Definition: The pineal gland is a small, secretory neuroendocrine organ located in the mid-line of the brain. Its primary function is to process information about the current light-dark cycle from the environment via the inner retina, and thereby synthesize melatonin during the dark portion of the cycle to help maintain normal sleep and circadian rhythms. Melatonin is suppressed by light to the retina and its secretion is controlled by the circadian timing system driven by the circadian pacemaker, the suprachiasmatic nucleus

Affected population: Older adults

Predisposing or influencing factors: No information available.

Fluoride theory and evidence: In 2006 in the USA, the National Research Council conducted a comprehensive review of the health effects of fluoride exposure. One conclusion reached was that fluoride is likely to impair pineal function and hence reduce melatonin production. The pineal gland is a small neuroendocrine organ situated near the centre of the brain. It sits outside of the blood-brain barrier, and thus the passage of fluoride is not restricted as it is in other areas of the brain. Its tissue is subject to mineralization, with calcification producing concretions or accumulations up to several millimetres in diameter. This calcification consists of hydroxyapatite, the same mineral found in bones and teeth. It has been found to accumulate high levels of fluoride even from low fluoride consumption due to fluoride's high affinity for hydroxyapatite. This vulnerability could increase the risk of fluoride poisoning in the pineal gland. In older individuals, fluoride measurements in the pineal gland have been shown to be roughly equivalent to those in teeth. Given fluoride's propensity to accumulate in the pineal gland, as well as the well-established associations between the degree of pineal gland calcification and human melatonin levels and disruption of numerous sleep-related outcomes (sleep efficiency, daytime tiredness, and sleep disturbance), further research is needed to examine the potential for fluoride to impact sleep outcomes.

Sources:

1. Aulinas A. Physiology of the Pineal Gland and Melatonin. In: Feingold KR, Anawalt B, Boyce A, Chrousos G, de Herder WW, Dhatariya K, *et al.*, editors. Endotext. South Dartmouth (MA): MDText.com, Inc.; 2000.
2. Cunningham JE, McCague H, Malin AJ, Flora D, Till C. Fluoride exposure and duration and quality of sleep in a Canadian population-based sample. Environmental Health. 2021 Dec;20(1):1-0.

Renal conditions

Renal calculi (Kidney stones)

Definition: Kidney stones (also called renal calculi, nephrolithiasis, or urolithiasis) are hard deposits made of minerals and salts that form inside your kidneys.

Affected population: The causes of kidney stones are eating calcium oxalate rich foods; purine rich foods; women with frequent urinary tract infections; and people with a genetic disorder cystinuria. Most ingested fluoride is excreted via the kidney, which is therefore exposed to relatively high fluoride concentrations. Kidney stones are most frequently diagnosed in people who drink less than 1 litre of water each day. Females and white people are more susceptible. Other risk factors include obesity, a diet with high levels of protein, salt, or glucose, hyperparathyroid condition, gastric bypass surgery, inflammatory bowel diseases that increase calcium absorption, taking medications such as triamterene diuretics, anti-seizure drugs, and calcium-based antacids.

Predisposing or influencing factors: Kidney stones are most frequently diagnosed in people who drink less than 1 litre of water each day. Females and white people are more susceptible. Other risk factors include obesity, a diet with high levels of protein, salt, or glucose, hyperparathyroid condition, gastric bypass surgery, inflammatory bowel diseases that increase calcium absorption, taking medications such as triamterene diuretics, anti-seizure drugs, and calcium-based antacids.

Fluoride theory and evidence: According to some studies, infants and children who are exposed to high doses of fluoride in their drinking water and diet have higher fluoride retention in their bodies, weakening their kidneys and making them more susceptible to kidney diseases later in life. Similarly, long-term fluoride exposure is associated with renal dysfunction in adults. However, due to the limited available research on the topic, there is no specific recommendations regarding fluoride intake and kidney disease.

Sources:

1. Public Health England. Water fluoridation: health monitoring report for England 2014.
2. Wu L, Fan C, Zhang Z, Zhang X, Lou Q, Guo N, Huang W, Zhang M, Yin F, Guan Z, Yang Y. Association between fluoride exposure and kidney function in adults: A cross-sectional study based on endemic fluorosis area in China. Ecotoxicology and Environmental Safety. 2021 Dec 1;225:112735.
3. Xiong X, Liu J, He W, Xia T, He P, Chen X, Yang K, Wang A. Dose–effect relationship between drinking water fluoride levels and damage to liver and kidney functions in children. Environmental research. 2007 Jan 1;103(1):112-6.

Birth or birthing abnormalities

Down syndrome

Definition: There are three types of Down syndrome: Trisomy 21 (95%): extra number 21 chromosomes in every cell in the body; Translocation (3–5%): an extra chromosome 21 is attached to another chromosome in every cell; Mosaic (1–2 %): mixture of cells, some with an extra chromosome 21 and some normal.

Affected population: Approximately 1% of all translocation trisomies are inherited. Although older moms have a greater likelihood of having a child with Down syndrome than younger mothers, children with Down syndrome are born to parents of all ages and socioeconomic backgrounds.

Predisposing or influencing factors: Down syndrome is associated with maternal age and genetic profile.

Fluoride theory and evidence: As fluoride can cross the placenta, the possibility of fluoride having a cytogenetic effect on the developing foetus is theoretically plausible.

Sources:

1. Public Health England. Water fluoridation: health monitoring report for England 2014.
2. Young N, Newton J, Morris J, Morris J, Langford J, Iloya J, Edwards D, Makhani S, Verne J. Community water fluoridation and health outcomes in England: a cross-sectional study. Community dentistry and oral epidemiology. 2015 Dec;43(6):550-9.

Congenital abnormalities

Definition: Congenital abnormalities are structural or functional anomalies develop during fetal life. Congenital abnormality include trisomy (trisomy 21, 13, and 18 only, ICD-9 codes 758.0, 758.1, 758.2), a neural tube defect (as defined by the EUROCAT system of classification, ICD-9 codes 740.0, 740.1, 740.2,

741.0, 741.9, and 742.0) or facial cleft (cleft palate, cleft lip with or without cleft palate, Pierre Robin syndrome, ICD-9 codes 749.0, 749.1,749.2, 756.03).

Affected population: Foetus and new-born

Predisposing or influencing factors: Most birth defects are caused by genetic or environmental factors or a combination of the two (multifactorial birth defects). In most cases, however, the cause is unknown. Some environmental factors include smoking, alcohol consumption, taking drugs of any type (apart for folic acid) that are not prescribed, not taking folic acid, uncontrolled diabetes, uncontrolled blood pressure, exposure to chemicals (lead, pesticides, and radiation), and specific infections (toxoplasmosis).

Fluoride theory and evidence: As fluoride can cross the placenta, the possibility of fluoride having a cytogenetic effect on the developing foetus is theoretically plausible.

Sources:

1. Lowry R, Steen N, Rankin J. Water fluoridation, stillbirths, and congenital abnormalities. Journal of Epidemiology & Community Health. 2003 Jul 1;57(7):499-500.
2. Public Health England. Water fluoridation: health monitoring report for England 2014.

Stillbirth

Definition: A stillbirth is a baby born dead after 24 weeks pregnancy. One in every 200 births ends in a stillbirth.

Affected population: Foetus

Predisposing or influencing factors: Pregnancy and labour complications including problems with the placenta, birth defects, infection, problems with the umbilical cord, high maternal blood pressure disorders, and other medical complications in the mother.

Fluoride theory and evidence: No information available.

Sources:

1. Lowry R, Steen N, Rankin J. Water fluoridation, stillbirths, and congenital abnormalities. Journal of Epidemiology & Community Health. 2003 Jul 1;57(7):499-500.

Preterm birth

Definition: Preterm birth occurs when a baby is born prematurely, before 37 weeks of pregnancy have been completed. The production of inflammatory mediators in the inflamed periodontal tissues could enter the bloodstream and trigger an inflammatory cascade in the uterus, leading to preterm birth.

Affected population: Foetus and neonate

Predisposing or influencing factors: Common causes of preterm birth include multiple pregnancies, infections, and chronic conditions such as diabetes and high blood pressure; however, the cause is often unknown. There could also be a genetic influence.

Fluoride theory and evidence: It has been proposed that the bacteria that cause dental caries and periodontitis can enter the bloodstream and trigger an inflammatory cascade, leading to preterm birth; fluoride, in reducing caries, can thus theoretically lower the risk of preterm birth However, there is limited evidence to support an association between periodontal disease and preterm birth, which have many more substantial risk factors in common, including smoking.

Sources:

1. Agueda A, Ramón JM, Manau C, Guerrero A, Echeverría JJ. Periodontal disease as a risk factor for adverse pregnancy outcomes: a prospective cohort study. Journal of clinical periodontology. 2008 Jan;35(1):16-22.
2. Heimonen A, Rintamäki H, Furuholm J, Janket SJ, Kaaja R, Meurman JH. Postpartum oral health parameters in women with preterm birth. Acta Odontologica Scandinavica. 2008 Jan 1;66(6):334-41.
3. Zhang X, Lu E, Stone SL, Diop H. Dental cleaning, community water fluoridation and preterm birth, Massachusetts: 2009–2016. Maternal and child health journal. 2019 Apr;23(4):451-8.

Infant abnormalities

Sudden infant death syndrome

Definition: Sudden infant death syndrome, sometimes referred to as cot death or crib death, is the sudden unexplained death of a child under the age of one year. Diagnosis requires that the death remain unexplained even after a thorough autopsy and detailed death scene investigation. Sudden infant death syndrome usually occurs during sleep.

Affected population: The risk for sudden infant death syndrome peaks between 2 and 3 months of age, and it occurs more often in male infants than in females.

Predisposing or influencing factors: Common factors that contribute to an increased risk are infant sleeping in the prone position (on their stomachs); use of soft bedding or unsafe beds (couches, daybeds, waterbeds); use of loose bedding materials such as blankets and pillows; overheating due to clothing, blankets or room temperature; mother's age younger than 20 years; mother smoking during pregnancy; mother receiving late or no prenatal care; premature birth or low birth weight; exposure to second-hand smoke

Fluoride theory and evidence: No information provided to explain a theoretical association

Sources:

1. Dick AE, Ford RP, Schluter PJ, Mitchell EA, Taylor BJ, Williams SM, Stewart AW, Becroft DM, Thompson JM, Scragg R, Hassall IB. Water fluoridation and the sudden infant death syndrome. The New Zealand Medical Journal. 1999 Aug 1;112(1093):286-9.

Appendix B PRISMA checklist

Table 43 Preferred Reporting Items for Systematic reviews and Meta-Analyses checklist

Topic	Item	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Executive summary
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Sections 1.1, 1.2
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Sections 1.2, 1.3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Section 2.3
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Sections 2.4.2.3, 2.4.2.4, 2.4.2.6, 2.4.2.7
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Appendix D
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Section 2.4.2.5
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Section 2.5

Topic	Item	Checklist item	Location where item is reported
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Section 2.5, Appendix F
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Section 2.5
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Section 2.6, Appendix G
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Appendix N
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Appendix H
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Section 2.5
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Section 2.7.2
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Section 2.7
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	n/a
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	n/a
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	n/a

Topic	Item	Checklist item	Location where item is reported
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Section 2.7.3
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Section 3.1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Appendix C
Study characteristics	17	Cite each included study and present its characteristics.	Section 3.2, Appendix L
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Appendix M
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Appendix N
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Sections 3.x.1.1 and 3.x.1.2 for each of eight outcomes
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	n/a
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	n/a
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	n/a
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	n/a
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Section 3.11

Topic	Item	Checklist item	Location where item is reported
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Section 4.14.2
	23b	Discuss any limitations of the evidence included in the review.	Sections 4.1.9, 4.2.9
	23c	Discuss any limitations of the review processes used.	Section 4.4
	23d	Discuss implications of the results for practice, policy, and future research.	Section 4.4
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Section 2.1
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Section 2.1
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	n/a
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	n/a
Competing interests	26	Declare any competing interests of review authors.	n/a
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	n/a

Source: Page *et al.* (2021)[28]

Appendix C Excluded studies

Table 44 Excluded studies with reasons for exclusion and data source

Reference	Reason for exclusion
Peer review literature/systematic reviews (excluded at full-text review)	
Adair SM, Hanes CM, Russell CM, Whitford GM. Dental caries and fluorosis among children in a rural Georgia area. Pediatric dentistry. 1999 Mar 1;21:81-5.	Exclude - Outcomes
Angelillo IF, Torre I, Nobile CG, Villari P. Caries and fluorosis prevalence in communities with different concentrations of fluoride in the water. Caries research. 1999;33(2):114-22.	Exclude - Outcomes
Archer NP, Napier TS, Villanacci JF. Fluoride exposure in public drinking water and childhood and adolescent osteosarcoma in Texas. Cancer Causes & Control. 2016 Jul;27(7):863-8.	Exclude - Intervention
Aung YM, Tin ST, Jelleyman T, Ameratunga S. Dental caries and previous hospitalisations among preschool children: Findings from a population-based study in New Zealand. NZ Med J. 2019 Apr 12;132(1493):44-53.	Exclude - Intervention
Azcurra AI, Battellino LJ, Calamari SE, de Cattoni ST, Kremer M, Lamberghini FC. Dental health status of students living in places supplied with drinking water of very high and very low levels of fluorides. Revista de saude publica. 1995 Oct 1;29(5):364-75.	Exclude - Outcomes
Bai R, Huang Y, Wang F, Guo J. Associations of fluoride exposure with sex steroid hormones among US children and adolescents, NHANES 2013–2016. Environmental Pollution. 2020 May 1;260:114003.	Exclude - Intervention
Booth JM, Mitropoulos CM, Worthington HV. A comparison between the dental health of 3-year-old children living in fluoridated Huddersfield and non-fluoridated Dewsbury in 1989. Community dental health. 1992 Jun 1;9(2):151-7.	Exclude - Outcomes
Brothwell DJ, Limeback H. Fluorosis risk in grade 2 students residing in a rural area with widely varying natural fluoride. Community dentistry and oral epidemiology. 1999 Apr;27(2):130-6.	Exclude - Intervention
Cauley JA, Murphy PA, Riley T, Black D. Public health bonus of water fluoridation: Does fluoridation prevent ostoporosis and its related fractures?. American Journal of Epidemiology 1991 Oct 1 (Vol. 134, No. 7, pp. 768-768).	Exclude - Study design
Chafe R, Aslanov R, Sarkar A, Gregory P, Comeau A, Newhook LA. Association of type 1 diabetes and concentrations of drinking water components in Newfoundland and Labrador, Canada. BMJ Open Diabetes Research and Care. 2018 Feb 1;6(1):e000466.	Exclude - Intervention
Chen BC. Epidemiological study on dental fluorosis and dental caries prevalence in communities with negligible, optimal, and above-optimal fluoride concentrations in drinking water supplies. Zhonghua ya yi xue hui za zhi. 1989 Sep 1;8(3):117-27.	Exclude - Intervention
Clark DC, Hann HJ, Williamson MF, Berkowitz J. Aesthetic concerns of children and parents in relation to different classifications of the Tooth Surface Index of Fluorosis. Community Dentistry and Oral Epidemiology. 1993 Dec;21(6):360-4.	Exclude - Outcomes

Reference	Reason for exclusion
Clarkson JJ, O'Mullane DM. Prevalence of enamel defects/fluorosis in fluoridated and non-fluoridated areas in Ireland. Community dentistry and oral epidemiology. 1992 Aug;20(4):196-9.	Exclude - Outcomes
Cochran JA, Ketley CE, Arnadóttir IB, Fernandes B, Koletsi-Kounari H, Oila A-M, *et al.* A comparison of the prevalence of fluorosis in 8-year-old children from seven European study sites using a standardized methodology. Community dentistry and oral epidemiology. 2004 Apr;32:28-33.	Exclude - Outcomes
Cooper C, Wickham C, Lacey RF, Barker DJ. Water fluoride concentration and fracture of the proximal femur. Journal of Epidemiology & Community Health. 1990 Mar 1;44(1):17-9.	Exclude - Intervention
Correia Sampaio F, Ramm von der Fehr F, Arneberg P, Petrucci Gigante D, Hatløy A. Dental Fluorosis and Nutritional Status of 6-to 11-Year-Old Children Living in Rural Areas of Paraíba, Brazil. Caries Research. 1999 Jan 1;33(1).	Exclude - Intervention
Downer MC, Blinkhorn AS, Holt RD, Wight C, Attwood D. Dental caries experience and defects of dental enamel among 12-year-old children in north London, Edinburgh, Glasgow and Dublin. Community Dentistry and Oral Epidemiology. 1994 Oct;22(5PT1):283-5.	Exclude - Outcomes
Ellwood RP, O'mullane DM. Dental enamel opacities in three groups with varying levels of fluoride in their drinking water. Caries research. 1995;29(2):137-42.	Exclude - Outcomes
Ellwood RP, O'Mullane D. The association between developmental enamel defects and caries in populations with and without fluoride in their drinking water. Journal of public health dentistry. 1996 Mar;56(2):76-80.	Exclude - Outcomes
Evans DJ, Rugg-Gunn AJ, Tabari ED, Butler T. The effect of fluoridation and social class on caries experience in 5-year-old Newcastle children in 1994 compared with results over the previous 18 years. Community dental health. 1996 Mar 1;13(1):5-10.	Exclude - Outcomes
Fabiani L, Leoni V, Vitali M, Parafati M, Rodolico SE, Cremisini C. Fluoride in water as a protective factor for bone fractures: preliminary data of an epidemiological study in Italy. IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences. 1995;233:135-40.	Exclude - Intervention
Feskanich D, Owusu W, Hunter DJ, Willett W, Ascherio A, Spiegelman D, *et al.* Use of toenail fluoride levels as an indicator for the risk of hip and forearm fractures in women. Epidemiology. 1998 Jul 1:412-6.	Exclude - Intervention
Fluegge K. Community water fluoridation predicts increase in age-adjusted incidence and prevalence of diabetes in 22 states from 2005 and 2010. Journal of water and health. 2016 Oct 1;14(5):864-77.	Exclude - Comparator
Forbes WF, Gentleman JF, Agwani N, Lessard S, McAiney CA. Geochemical risk factors for mental functioning, based on the Ontario Longitudinal Study of Aging (LSA) VI. The effects of iron on the associations of aluminum and fluoride water concentrations and of pH with mental functioning, based on results obtained from the LSA and from death certificates mentioning dementia. Canadian Journal on Aging/La Revue canadienne du vieillissement. 1997 Jan;16(1):142-59.	Exclude - Intervention
Freni SC, Gaylor DW. International trends in the incidence of bone cancer are not related to drinking water fluoridation. Cancer. 1992 Aug 1;70(3):611-8.	Exclude - Comparator

Reference	Reason for exclusion
Gelberg KH, Fitzgerald EF, Hwang SA, Dubrow R. Fluoride exposure and childhood osteosarcoma: a case-control study. American Journal of Public Health. 1995 Dec;85(12):1678-83.	Exclude - Intervention
Griffin SO, Beltrán ED, Lockwood SA, Barker LK. Esthetically objectionable fluorosis attributable to water fluoridation. Community dentistry and oral epidemiology. 2002 Jun;30(3):199-209.	Exclude - Outcomes
Grimaldo M, Borjaaburto VH, Ramirez AL, Ponce M, Rosas M, Diazbarriga F. Endemic fluorosis in San-Luis-Potosi, Mexico. 1. Identification of risk-factors associated with human exposure to fluoride. Environmental Research. 1995 Jan 1;68(1):25-30.	Exclude - Intervention
Harding MA, Whelton H, O Mullane DM, Cronin M, Warren JJ. Primary tooth fluorosis in 5-year-old schoolchildren in Ireland. European journal of paediatric dentistry. 2005 Sep 1;6(3):155.	Exclude - Outcomes
Hawew RM, Ellwood RP, Hawley GM, Worthington HV, Blinkhorn AS. Dental caries in children from two Libyan cities with different levels of fluoride in their drinking water. Community dental health. 1996 Sep 1;13(3):175-7.	Exclude - Intervention
Heintze SD, Bastos JR, Bastos R. Urinary fluoride levels and prevalence of dental fluorosis in three Brazilian cities with different fluoride concentrations in the drinking water. Community Dentistry and Oral Epidemiology. 1998 Oct;26(5):316-23.	Exclude - Outcomes
Heller KE, Eklund SA, Burt BA. Dental caries and dental fluorosis at varying water fluoride concentrations. Journal of Public Health Dentistry. 1997 Sep;57(3):136-43.	Exclude - Outcomes
Helte E, Donat Vargas C, Kippler M, Wolk A, Michaëlsson K, Åkesson A. Fluoride in Drinking Water, Diet, and Urine in Relation to Bone Mineral Density and Fracture Incidence in Postmenopausal Women. Environmental health perspectives. 2021 Apr 6;129(4):047005.	Exclude - Intervention
Hillier S, Cooper C, Kellingray S, Russell G, Hughes H, Coggon D. Fluoride in drinking water and risk of hip fracture in the UK: a case-control study. The Lancet. 2000 Jan 22;355(9200):265-9.	Exclude - Intervention
Hillier S, Kellingray S, Coggon D, Inskip H, Wallace I, Russell RG, *et al.* Water fluoridation and fracture of the proximal femur. Journal of Bone and Mineral Research. 1997 Sep 5;12(9).	Exclude - Study design
Hong CY, Hong YC, Guo MK, Hsieh CC, Chen RS. Prevalence of mottled enamel after 12 years of water fluoridation in Chung-hsing New Village. Journal of the Formosan Medical Association. 1990 Mar 1;89(3):225-30.	Exclude - Outcomes
Hoover RN, Devesa SS, Cantor KP, Lubin JH, Fraumeni JF. Fluoridation of drinking water and subsequent cancer incidence and mortality. Appendix E in: Review of Fluoride Benefits and Risks. Report of the Ad Hoc Subcommittee on Fluoride of the Committee to Coordinate Environmental Health and Related Programs. US Public Health Service. 1991 Feb.	Exclude - Incomplete
Ibrahim YE, Abuaffan AH, Bjorvatn K. Prevalence of dental fluorosis in Sudanese children from two villages with 0.25 and 2.56 ppm fluoride in the drinking water. International journal of paediatric dentistry. 1995 Dec;5(4):223-9.	Exclude - Intervention
Ismail AI, Brodeur JM, Kavanagh ME, Boisclair G, Tessier C, Picotte L. Prevalence of dental caries and dental fluorosis in students, 11–17 years of age, in fluoridated and non-fluoridated cities in Quebec. Caries research. 1990;24(4):290-7.	Exclude - Outcomes

Reference	Reason for exclusion
Jackson RD, Kelly SA, Katz B, Brizcndine E, Stookey GK. Dental fluorosis in children residing in communities with different water fluoride levels: 33-month follow-up. Pediatric dentistry. 1999 Jul 1;21:248-54.	Exclude - Outcomes
Jacqmin H, Commenges D, Letenneur L, Barberger-Gateau P, Dartigues JF. Components of drinking water and risk of cognitive impairment in the elderly. American journal of epidemiology. 1994 Jan 1;139(1):48-57.	Exclude - Intervention
Jacqmin-Gadda H, Commenges D, Dartigues JF. Fluorine concentration in drinking water and fractures in the elderly. JAMA. 1995 Mar 8;273(10):775-6.	Exclude - Intervention
Jacqmin-Gadda H, Fourrier A, Commenges D, Dartigues JF. Risk factors for fractures in the elderly. Epidemiology. 1998 Jul 1:417-23.	Exclude - Intervention
Johnson Jr J. Water fluoridation. Today's FDA: official monthly journal of the Florida Dental Association. 2014;26(5):32-3.	Exclude - Study design
Jones C, Taylor G, Woods K, Whittle G, Evans D, Young P. Jarman underprivileged area scores, tooth decay and the effect of water fluoridation. Community dental health. 1997 Sep 1;14(3):156-60.	Exclude - Outcomes
Jones CM, Worthington H. The relationship between water fluoridation and socioeconomic deprivation on tooth decay in 5-year-old children. British dental journal. 1999 Apr;186(8):397-400.	Exclude - Outcomes
Jones CM, Worthington H. Water fluoridation, poverty and tooth decay in 12-year-old children. Journal of Dentistry. 2000 Aug 1;28(6):389-93.	Exclude - Outcomes
Jooste PL, Weight MJ, Kriek JA, Louw AJ. Endemic goitre in the absence of iodine deficiency in schoolchildren of the Northern Cape Province of South Africa. European journal of clinical nutrition. 1999 Jan;53(1):8-12.	Exclude - Intervention
Kalsbeek H, Kwant GW, Groeneveld A, Dirks OB, Van Eck AA, Theuns HM. Caries experience of 15-year-old children in The Netherlands after discontinuation of water fluoridation. Caries research. 1993;27(3):201-5.	Exclude - Outcomes
Karagas MR, Baron JA, Barrett JA, Jacobsen SJ. Patterns of fracture among the United States elderly: geographic and fluoride effects. Annals of epidemiology. 1996 May 1;6(3):209-16.	Exclude - Intervention
Kelman AM. Fluoridation--the Israel experience. Community dental health. 1996 Sep 1;13:42-6.	Exclude - Outcomes
Kim FM, Hayes C, Williams PL, Whitford GM, Joshipura KJ, Hoover RN, *et al.* National Osteosarcoma Etiology Group. An assessment of bone fluoride and osteosarcoma. Journal of dental research. 2011 Oct;90(10):1171-6.	Exclude - Intervention
Kumar JV, Swango PA. Fluoride exposure and dental fluorosis in Newburgh and Kingston, New York: policy implications. Community dentistry and oral epidemiology. 1999 Jun;27(3):171-80.	Exclude - Outcomes
Kunzel W, Fischer T. Rise and fall of caries prevalence in German towns with different F concentrations in drinking water. Caries research. 1997 May 1;31(3):166.	Exclude - Outcomes
Kurttio P, Gusta vs. son N, Vartiainen T, Pekkanen J. Exposure to natural fluoride in well water and hip fracture: a cohort analysis in Finland. American journal of epidemiology. 1999 Oct 15;150(8):817-24.	Exclude - Intervention
Kurttio P, Gusta vs. son N, Vartiainen T, Pekkanen J. Exposure to natural fluoride in well water and hip fracture: a cohort analysis in Finland. American journal of epidemiology. 1999 Oct 15;150(8):817-24.	Exclude - Intervention
Lan CF, Lin IF, Wang SJ. Fluoride in drinking water and the bone mineral density of women in Taiwan. International journal of epidemiology. 1995 Dec 1;24(6):1182-7.	Exclude - Intervention

Reference	Reason for exclusion
Levy M, Leclerc BS. Fluoride in drinking water and osteosarcoma incidence rates in the continental United States among children and adolescents. Cancer epidemiology. 2012 Apr 1;36(2):e83-8.	Exclude - Intervention
Levy SM, Eichenberger-Gilmore JM, Warren JJ, Kavand G, Letuchy E, Broffitt B, *et al.* Associations of fluoride intake with children's cortical bone mineral and strength measures at age 11. Journal of public health dentistry. 2018 Sep;78(4):352-9.	Exclude - Intervention
Levy SM, Warren JJ, Phipps K, Letuchy E, Broffitt B, Eichenberger-Gilmore J, *et al.* Effects of life-long fluoride intake on bone measures of adolescents: a prospective cohort study. Journal of dental research. 2014 Apr;93(4):353-9.	Exclude - Intervention
Li Y, Liang C, Slemenda CW, Ji R, Sun S, Cao J, *et al.* Effect of long-term exposure to fluoride in drinking water on risks of bone fractures. J Bone Miner Res Off J Am Soc Bone Miner Res. 2001 May;16(5):932–9.	Exclude - Intervention
Lin. The relationship of a low-iodine an high-fluoride environment to subclinical cretinism in Xinjiang. Xinjiang Institute for Endemic Disease Control and Research, Office of Leading Group for Endemic Disease Control of Hetian Prefectural Committee of the Communist Party of China and County Health and Endemic Prevention Station, Yutian, Xinjiang. 1991. Unpublished report submitted through NHS CRD web site.	Exclude - Intervention
Loh T. Thirty-eight years of water fluoridation--the Singapore scenario. Community Dental Health. 1996 Sep 1;13:47-50.	Exclude - Outcomes
Malin AJ, Lesseur C, Busgang SA, Curtin P, Wright RO, Sanders AP. Fluoride exposure and kidney and liver function among adolescents in the United States: NHANES, 2013–2016. Environment international. 2019 Nov 1;132:105012.	Exclude - Intervention
Malin AJ, Till C. Exposure to fluoridated water and attention deficit hyperactivity disorder prevalence among children and adolescents in the United States: an ecological association. Environmental Health. 2015 Dec;14(1):1-0.	Exclude - Comparator
Masztalerz A, Masztalerzowa Z, Szymańska M, Tomelka J. Fluorine and the dentition. Fortschritte der Kieferorthopadie. 1990 Aug 1;51(4):234-7.	Exclude - Intervention
McGrady MG, Ellwood RP, Maguire A, Goodwin M, Boothman N, Pretty IA. The association between social deprivation and the prevalence and severity of dental caries and fluorosis in populations with and without water fluoridation. BMC public health. 2012 Dec;12(1):1-7.	Exclude - Outcomes
Milsom K, Mitropoulos CM. Enamel defects in 8-year-old children in fluoridated and non-fluoridated parts of Cheshire. Caries research. 1990;24(4):286-9.	Exclude - Outcomes
Moss ME, Kanarek MS, Anderson HA, Hanrahan LP, Remington PL. Osteosarcoma, seasonality, and environmental factors in Wisconsin, 1979–1989. Archives of Environmental Health: An International Journal. 1995 Jun 1;50(3):235-41.	Exclude - Intervention
Murray JJ, Breckon JA, Reynolds PJ, Tabari ED, Nunn JH. The effect of residence and social class on dental caries experience in 15-16-year-old children living in three towns (natural fluoride, adjusted fluoride and low fluoride) in the north east of England. British dental journal. 1991 Nov;171(10):319-22.	Exclude - Outcomes

Reference	Reason for exclusion
Nahum LH. Mutual Medical Dental Problems: Fluoridation of Water Supply. 1965. Connecticut medicine. 2015 Mar 1;79(3):177-9.	Exclude - Study design
Näsman P, Ekstrand J, Granath F, Ekbom A, Fored CM. Estimated drinking water fluoride exposure and risk of hip fracture: a cohort study. Journal of dental research. 2013 Nov;92(11):1029-34.	Exclude - Intervention
Nunn JH, Murray JJ, Reynolds P, Tabari D, Breckon J. The prevalence of developmental defects of enamel in 15-16-year-old children residing in three districts (natural fluoride, adjusted fluoride, low fluoride) in the north east of England. Community dental health. 1992 Sep 1;9(3):235-47.	Exclude - Outcomes
Nunn JH, Rugg-Gunn AJ, Ekanayake L, Saparamadu KD. Prevalence of developmental defects of enamel in areas with differing water fluoride levels and socio-economic groups in Sri Lanka and England. International dental journal. 1994 Apr 1;44(2):165-73.	Exclude - Outcomes
Oweis R, Levy S, Warren J, Gilmore JE, Burns T, Saha P, *et al.* Associations of Fluoride Intake with Adolescents' pQCT-derived Bone Outcome Measures at Age 17. In Washington State Convention Center; 2015. Available from: https://www.asbmr.org/education/AbstractDetail?aid=0b402d4b-d06e-4dc3-923d-49b15d6dd946	Exclude - Study design
Oweis RR, Levy SM, Eichenberger-Gilmore JM, Warren JJ, Burns TL, Janz KF, *et al.* Fluoride intake and cortical and trabecular bone characteristics in adolescents at age 17: A prospective cohort study. Community dentistry and oral epidemiology. 2018 Dec;46(6):527-34.	Exclude - Intervention
Patient's page. Water fluoride. J - Okla Dent Assoc. 2014 Feb;105(2):8	Exclude - Study design
Perrott KW. Fluoridation and attention deficit hyperactivity disorder–a critique of Malin and Till (2015). British dental journal. 2017 Dec;223(11):819-22.	Exclude - Comparator
Phipps KR, Burt BA. Water-borne fluoride and cortical bone mass: a comparison of two communities. Journal of dental research. 1990 Jun;69(6):1256-60.	Exclude - Comparator
Phipps KR, Orwoll ES, Bevan L. The association between water-borne fluoride and bone mineral density in older adults. Journal of dental research. 1998 Sep;77(9):1739-48.	Exclude - Intervention
Provart SJ, Carmichael CL. The relationship between caries, fluoridation and material deprivation in five-year-old children in Country Durham. Community dental health. 1995 Dec 1;12(4):200-3.	Exclude - Outcomes
Public Health England. Water fluoridation: health monitoring report for England 2014.	Exclude - Duplicate data
Riley JC, Lennon MA, Ellwood RP. The effect of water fluoridation and social inequalities on dental caries in 5-year-old children. International Journal of Epidemiology. 1999 Apr 1;28(2):300-5.	Exclude - Outcomes
Riordan PJ, Banks JA. Dental fluorosis and fluoride exposure in Western Australia. Journal of dental research. 1991 Jul;70(7):1022-8.	Exclude - Outcomes
Rugg-Gunn AJ, Al-Mohammadi SM, Butler TJ. Effects of fluoride level in drinking water, nutritional status, and socio-economic status on the prevalence of developmental defects of dental enamel in permanent teeth in Saudi 14-year-old boys. Caries research. 1997;31(4):259-67.	Exclude - Outcomes
Rwenyonyi MC, Birkeland JM, Bjorvatn K, Haugejorden O. Dental fluorosis in Ugandans related to fluoride in drinking water and altitude. InJournal of Dental Research 1998 Jan 1 (Vol. 77, pp. 794-794).	Exclude - Intervention

Reference	Reason for exclusion
Schwartz GG. Eye cancer incidence in US states and access to fluoridated water. Cancer Epidemiology and Prevention Biomarkers. 2014 Sep 1;23(9):1707-11.	Exclude - Intervention
Selwitz RH, Nowjack-Raymer RE, Kingman A, Driscoll WS. Dental caries and dental fluorosis among schoolchildren who were lifelong residents of communities having either low or optimal levels of fluoride in drinking water. Journal of public health dentistry. 1998 Mar;58(1):28-35.	Exclude - Outcomes
Selwitz RH, Nowjack-Raymer RE, Kingman A, Driscoll WS. Prevalence of dental caries and dental fluorosis in areas with optimal and above-optimal water fluoride concentrations: a 10-year follow-up survey. Journal of public health dentistry. 1995 Mar;55(2):85-93.	Exclude - Outcomes
Seppä L, Kärkkäinen S, Hausen H. Caries frequency in permanent teeth before and after discontinuation of water fluoridation in Kuopio, Finland. Community dentistry and oral epidemiology. 1998 Aug;26(4):256-62.	Exclude - Outcomes
Seppä L, Kärkkäinen S, Hausen H. Caries Trends 1992–1998 in Two Low–Fluoride Finnish Towns Formerly with and without Fluoridation. Caries research. 2000;34(6):462-8.	Exclude - Outcomes
Singh PP, Dhing S, Bhatnagar R, Kothari S, Dhar V. Evidence suggesting that high intake of fluoride provokes nephrolithiasis in tribal populations. Urological research. 2001 Aug;29(4):238-44.	Exclude - Intervention
Skotowski MC, Hunt RJ, Levy SM. Risk factors for dental fluorosis in pediatric dental patients. Journal of Public Health Dentistry. 1995 Jun;55(3):154-9.	Exclude - Outcomes
Slade GD, Spencer AJ, Davies MJ, Stewart JF. Caries experience among children in fluoridated Townsville and unfluoridated Brisbane. Australian and New Zealand journal of public health. 1996 Dec;20(6):623-9.	Exclude - Outcomes
Sowers MR, Clerk KM, Jannausch ML, Wallace RB. A prospective study of bone mineral content and fracture in communities with differential fluoride exposure. American journal of epidemiology. 1991 Apr 1;133(7):649-60.	Exclude - Comparator
Steiner GG. Cancer incidence rates and environmental factors: an ecological study. Journal of environmental pathology, toxicology and oncology. 2002;21(3).	Exclude - Intervention
Stephen KW, Macpherson LM, Gorzo I, Gilmour WH. Effect of fluoridated salt intake in infancy: a blind caries and fluorosis study in 8th grade Hungarian pupils. Community dentistry and oral epidemiology. 1999 Jun;27(3):210-5.	Exclude - Intervention
Tabari ED, Ellwood R, Rugg-Gunn AJ, Evans DJ, Davies RM. Dental fluorosis in permanent incisor teeth in relation to water fluoridation, social deprivation and toothpaste use in infancy. British dental journal. 2000 Aug;189(4):216-20.	Exclude - Outcomes
Takahashi K, Akiniwa K, Narita K. Regression analysis of cancer incidence rates and water fluoride in the USA based on IACR/IARC (WHO) data (1978-1992). Journal of Epidemiology. 2001;11(4):170-9.	Exclude - Intervention
Talpos Sara. They persisted. Science. 2019 May 17;364(6441):622–6.	Exclude - Intervention
Vignarajah S. Dental caries experience and enamel opacities in children residing in urban and rural areas of Antigua with different levels of natural fluoride in drinking water. Community dental health. 1993 Jun 1;10(2):159-66.	Exclude - Intervention
Villa AE, Guerrero S, Villalobos J. Estimation of optimal concentration of fluoride in drinking water under conditions prevailing in Chile. Community dentistry and oral epidemiology. 1998 Aug;26(4):249-55.	Exclude - Outcomes

Reference	Reason for exclusion
Wang XC, Kawahara K, Guo XJ. Fluoride contamination of groundwater and its impacts on human health in Inner Mongolia area. Journal of Water Supply: Research and Technology—AQUA. 1999 Jun;48(4):146-53.	Exclude - Intervention
Warnakulasuriya KA, Balasuriya S, Perera PA, Peiris LC. Determining optimal levels of fluoride in drinking water for hot, dry climates-a case study in Sri Lanka. Community dentistry and oral epidemiology. 1992 Dec;20(6):364-7.	Exclude - Outcomes
Weerheijm KL, Kidd EA, Groen HJ. The effect of fluoridation on the occurrence of hidden caries in clinically sound occlusal surfaces. Caries research. 1997;31(1):30-4.	Exclude - Outcomes
Whelton H, Crowley E, O'Mullane D, Donaldson M, Kelleher V, Cronin M. Dental caries and enamel fluorosis among the fluoridated and non-fluoridated populations in the Republic of Ireland in 2002. Community dental health. 2004 Mar 1;21(1):37-44.	Exclude - Outcomes
Yiamouyiannis JA. Fluoridation and cancer. The biology and epidemiology of bone and oral cancer related to fluoridation. Fluoride. 1993;26(2):83-96.	Exclude - Study design
Zhao LB, Liang GH, Zhang DN, Wu XR. Effect of a high fluoride water supply on children's intelligence. Fluoride. 1996 Nov 1;29(4):190-2.	Exclude - Intervention
Grey literature search	
Abrams S, Beltrán-Aguilar E, Martinez-Mier EA, Kumar J, Slade GD, Gooch B. Water fluoridation: safety, effectiveness and value in oral health: a symposium at the 2014 annual meeting of the American and Canadian associations for dental research. J. Can. Dent. Assoc. 2015 Jan 1;80:f16.	Exclude - Study design
Canadian Dental Association Magazine. 2014. Trois-Rivières to reintroduce community water fluoridation. Available from: https://www.cda-adc.ca/en/services/essentials/2014/issue7/index.html#17	Exclude - Study design
Dental Council of Ireland. 2019. Dental Health Foundation [overview of function]. . Available from: http://www.dentalcouncil.ie/	Exclude - Study design
Department of Health. Ministers for Health and Social Protection publish Smile agus Sláinte – the National Oral Health Policy [Internet]. 2019. Available from: https://www.gov.ie/en/press-release/09fbb1-ministers-for-health-and-social-protection-publish-smile-agus-slaint/	Exclude - Study design
Department of Housing, Local Government and Heritage. Drinking water quality [Internet]. 2016. Available from: https://www.gov.ie/en/publication/3870f-drinking-water-quality/	Exclude - Study design
Department of Minister for Housing, Planning; Local Government. 2017. EUROPEAN UNION (DRINKING WATER) (AMENDMENT) REGULATIONS 2017 [S.I. No. 464 of 2017]	Exclude - Study design
Garvin J. National Academies 'strongly recommends' third revision to fluoride monograph [Internet]. American Dental Association News. 2021. Available from: https://www.ada.org/en/publications/ada-news/2021-archive/february/national-academies-strongly-recommends-third-revision-to-fluoride-monograph	Exclude - Study design
Irish Expert Body on Fluorides and Health. 2020. Minutes of proceedings of the Plenary Meeting (Zoom) of the Irish Expert Body on Fluorides and Health held on Monday 14th December 2020 @3pm. Available from: https://www.fluoridesandhealth.ie/assets/files/pdf/plenary_minutes_14th_december_2020pdf_final.pdf	Exclude - Study design

Reference	Reason for exclusion
Jack B, Ayson M, Lewis S, Irving A, Agresta B, Ko H, *et al.* Health effects of water fluoridation: evidence evaluation report. Canberra (Australia): National Health and Medical Council. 2016.	Exclude - Study design
Jack, B, Ayson, M, Lewis, S, Irving, A, Agresta, B, Ko, H, *et al.* Health Effects of Water Fluoridation: Technical Report, report to the National Health and Medical Research Council, Canberra. 2016.	Exclude - Study design
Maybury C, Jacob M, Flanders JM, Horowitz AM. Seeking community water fluoridation information on state health department websites. Plos one. 2021 May 20;16(5):e0251139.	Exclude - Study design
Ministry of Health, New Zealand (2015). Rural Agricultural Drinking-water Supply Guideline. Available from https://www.health.govt.nz/publication/rural-agricultural-drinking-water-supply-guideline.	Exclude - Study design
Ministry of Health, New Zealand. 2019. Revised Drinking-Water Safety Plan Guidance Material – "Handbook for Preparing a Water Safety Plan". Available from: https://www.waternz.org.nz/Story?Action=View&Story_id=1023	Exclude - Study design
Ministry of Health, New Zealand. Drinking-water Standards for New Zealand 2005 (Revised 2018). Ministry of Health, New Zealand (2018). Available from: https://www.health.govt.nz/publication/drinking-water-standards-new-zealand-2005-revised-2018	Exclude - Study design
Ministry of Health, New ZealandGuidelines for Drinking-water Quality Management for New Zealand [Internet]. Ministry of Health, New Zealand. Available from: https://www.health.govt.nz/publication/guidelines-drinking-water-quality-management-new-zealand	Exclude - Study design
Montgomery T. 2021. [Proposed Postgraduate Research Opportunity] An investigation into fluoride exposure in women of child-bearing age in water-fluoridated versus non fluoridated areas. Athlone Institute of Technology. Available from: https://www.ait.ie/uploads/downloads/An_investigation_into_fluoride_exposure_in_women1.pdf	Exclude - Study design
Moore D, Poynton M. Review of the benefits and costs of water fluoridation in New Zealand. Sapere Research Group; 2015 Sep.	Exclude - Study design
National Academies of Sciences, Engineering, and Medicine. 2021. Quality Water from Every Tap: Proceedings of a Workshop–in Brief. Washington, DC: The National Academies Press Available from: https://www.nap.edu/catalog/26069/quality-water-from-every-tap-proceedings-of-a-workshop-in	Exclude - Study design
National Academies of Sciences, Engineering, and Medicine. 2021. Review of the Revised NTP Monograph on the Systematic Review of Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Letter Report. Washington, DC: The National Academies Press.https://doi.org/10.17226/26030.	Exclude - Study design
National Health and Medical Research Council (NHMRC) 2017, Information paper – Water fluoridation: dental and other human health outcomes, report prepared by the Clinical Trials Centre at University of Sydney, NHMRC; Canberra.	Exclude - Study design
National Health and Medical Research Council. 2017. Water Fluoridation and Human Health in Australia: Questions and Answers. Available from: https://www.nhmrc.gov.au/sites/default/files/documents/attachments/water-fluoridationqa.pdf	Exclude - Study design

Reference	Reason for exclusion
NHMRC Public Statement 2017: Water Fluoridation and Human Health in Australia. Available from: https://www.nhmrc.gov.au/sites/default/files/documents/reports/fluoridation-public-statement.pdf	Exclude - Study design
National Health and Medical Research Council. 2017. New Zealand Drinking-water Safety Plan Framework [Internet]. Ministry of Health NZ. Available from: https://www.health.govt.nz/publication/new-zealand-drinking-water-safety-plan-framework	Exclude - Study design
O Mullane DM, Baez RJ, Jones S, Lennon MA, Petersen PE, Rugg-Gunn AJ, *et al.* Fluoride and oral health. Community dental health. 2016 Jun 1;33(2):69-99.	Exclude - Study design
Pullishery F, Panchmal GS, Siddique S, Palliyal S. Status of Water Fluoridation Status of Water Fluoridation-An Update From the Asian Countries the Asian Countries.	Exclude - Study design
Reilly A. Fluoridation of water: a literature review of risks or benefits for the population in Ireland exposed to the current levels–including a European policy examination of water fluoridation practices.	Exclude - Study design
Mariño R, Zaror C. Economic evaluations in water-fluoridation: a scoping review. BMC oral health. 2020 Dec;20(1):1-2.	Exclude - Study design
Slade GD, Grider WB, Maas WR, Sanders AE. Water fluoridation and dental caries in US children and adolescents. Journal of dental research. 2018 Sep;97(10):1122-8.	Exclude - Study design
The Irish Expert Body on Fluorides and Health. Code of Practice on the Fluoridation of Drinking Water 2016 [Internet]. 2016. Available from: https://www.fluoridesandhealth.ie/resources/code-of-practice-on-the-fluoridation-of-drinking-water-2016/	Exclude - Study design
United States Environmental Protection Agency. 2018.	Exclude - Study design
World Health Organization. 2019. Preventing Disease Through Healthy Environments: Inadequate or Excess Fluoride: A Major Public Health Concern. Available from: https://apps.who.int/iris/rest/bitstreams/1257901/retrieve	Exclude - Study design
Reference/citation chasing	
Green R, Till C, Cantoral A, Lanphear B, Martinez-Mier E, Ayotte P, *et al.* Associations between urinary, dietary, and water fluoride concentrations among children in Mexico and Canada. Toxics. 2020 Dec;8(4):110.	Exclude - Outcomes
Han DH, Sun BC, Lim SY, Kim HD and Paek D. (2011) Association of fluoride exposure and bone mineral density: A comparison of area with individual (Conference Paper Abstract from the 22nd Annual Conference of the International Society for Environmental Epidemiology, ISEE 2010). Epidemiology, 22(1): S240-S241.	Exclude - Language
Lamberg M, Hausen H, Vartiainen T. Symptoms experienced during periods of actual and supposed water fluoridation. Community dentistry and oral epidemiology. 1997 Aug;25(4):291-5.	Exclude - Outcomes
Nicole W. Denser but Not Stronger? Fluoride-Induced Bone Growth and Increased Risk of Hip Fractures. Environmental Health Perspectives. 2021 Jul 12;129(7):074001.	Exclude - Study design
Pearce EN. Is fluoridated drinking water associated with increased hypothyroidism risk?. Clinical Thyroidology. 2015 Apr 1;27(4):100-1.	Exclude - Study design

Reference	Reason for exclusion
Riddell JK, Malin AJ, McCague H, Flora DB, Till C. Urinary Fluoride Levels among Canadians with and without Community Water Fluoridation. International Journal of Environmental Research and Public Health. 2021 Jan;18(12):6203.	Exclude - Outcomes
Sohn W, Heller KE, Burt BA. Fluid consumption related to climate among children in the United States. Journal of public health dentistry. 2001 Jun;61(2):99-106.	Exclude - Outcomes
Sowers M, Whitford GM, Clark MK, Jannausch ML. Elevated serum fluoride concentrations in women are not related to fractures and bone mineral density. The journal of nutrition. 2005 Sep 1;135(9):2247-52.	Exclude - Intervention
Sowers MR, Clerk KM, Jannausch ML, Wallace RB. A prospective study of bone mineral content and fracture in communities with differential fluoride exposure. American journal of epidemiology. 1991 Apr 1;133(7):649-60.	Exclude - Intervention
Strunecka A, Strunecky O. Chronic fluoride exposure and the risk of autism spectrum disorder. International journal of environmental research and public health. 2019 Jan;16(18):3431.	Exclude - Study design
Takahashi K, Akiniwa K, Narita K. Regression analysis of cancer incidence rates and water fluoride in the USA based on ICAR/IARC (WHO) data (1978–1992). International agency for research in cancer. J Epidemial. 2001;11:170-9.	Exclude - Intervention
Uyghurturk DA, Goin DE, Martinez-Mier EA, Woodruff TJ, DenBesten PK. Maternal and fetal exposures to fluoride during mid-gestation among pregnant women in northern California. Environmental Health. 2020 Dec;19(1):1-9.	Exclude - Outcomes
Waugh DT, Potter W, Limeback H, Godfrey M. Risk assessment of fluoride intake from tea in the Republic of Ireland and its implications for public health and water fluoridation. International journal of environmental research and public health. 2016 Mar;13(3):259.	Exclude - Intervention

Appendix D Search strategies

Table 45 Search strategy for Ovid MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions ® 1946 to 19 May 2021

Line number	Search terms	Number of results
1	(fluorid* or fluorin* or flurid* or florin*).mp./	101456
2	(fluorid* or fluorin* or flurid* or florin*).sh.	48551
3	exp Fluoride/ or exp Fluorides/ or exp Fluoridation/ or exp Fluorine/	48393
4	or/1-3.	103879
5	exp Water/ or exp Water Supply/	1022317
6	water.mp.	63861
7	or/5-6	1022317
8	blind*.mp.	381951
9	Case-Control Studies/	300970
10	case-control.mp.	349665
11	("case series" or "time series" or "before and after").mp.	403958
12	Cohort Studies/	284306
13	cohort analysis/	284306
14	cohort*.mp.	769380
15	control*.mp.	3551
16	clinical trial/	528796
17	Cohort Studies/	284306
18	cross over.mp.	50204
19	crossover.mp.	64120
20	crossover.mp.	66092
21	cross sectional.mp.	526686
22	double-blind method/	164386
23	doubl*.mp.	661351
24	exp clinical trial/	893059
25	comparative study/	1889651
26	comparative stud*.mp.	1945428
27	controlled clinical trial/	94165
28	Correlation study/	1891
29	cross sectional study/	366481
30	Ecological study/ or ecological stud*.mp. or ecological study.mp.	5639
31	evaluation study/	257887
32	evaluation stud*.mp.	384915
33	Follow-Up Studies/ or followup.mp. or follow-up.mp.	1416382
34	(health* adj2 survey*).mp.	138054
35	Incidence/	275372
36	incidence.mp.	902477
37	mask*.mp.	92701
38	Mortality/	46640
39	mortality.tw.	815698
40	observational study/	99007

Line number	Search terms	Number of results
41	Placebos/	35518
42	placebo*.mp.	239647
43	predict*.mp.	1795120
44	Prevalence/	308006
45	prevalence.mp.	759067
46	prognos*.mp	932742
47	random*.mp.	1462487
48	Randomized Controlled Trial/	530996
49	random allocation/	105347
50	risk.mp.	2735214
51	exp Research Design/	460112
52	Single-Blind Method/	30191
53	singl*.mp.	1893420
54	trebl*.mp.	496
55	tripl*.mp.	138238
56	volunteer*.mp	219353
57	or/8-56	13825297
58	4 and 7 and 57	7659
59	58 not (exp animals/ not humans.sh.)	6653
60	limit 59 to yr="2014 -Current"	2180
61	60 not (letter or comment or editorial or newspaper article).pt.	2154

Search date: 19 May 2021

Table 46 Search strategy for Ovid Embase 1974 to 19 May 2021

Line number	Search terms	Number of results
1	Fluoridation/	2335
2	Fluorine/	9248
3	exp Fluoride/	19843
4	(fluorid$ or fluorin$ or flurin$ or flurid$).ti,ab.	57481
5	or/1-4	65402
6	exp water supply/ or ground water/ or water analysis/ or water management/ or water standard/ or water absorption/ or exp tap water/ or exp water/	473368
7	exp Case Control Study/ or exp Controlled Study/ or exp Major Clinical Study/	9017332
8	("case series" or "time series" or "before and after").mp.	504515
9	exp Clinical Trial/	1451591
10	exp Cohort Studies/ or exp Cohort Analysis/ or cohort*.mp.	1236108
11	exp Correlation study/ or correlation coefficient.mp. or data correlation.mp. or correlation analysis.mp. or cross correlation.mp. or correlation.mp. or correlation function.mp.	1295581
12	Cross-Over Studies/ or (crossover or cross over or cross-over).mp.	101949
13	exp cross sectional study/ or (cross sectional or cross-sectional).mp.	597703
14	double-blind method/ or doubl*.mp. or blind.mp.	774717

Line number	Search terms	Number of results
15	exp ecological validity/	725
16	exp evaluation study/	74971
17	exp evaluation research/	2013
18	exp follow up/	1614990
19	exp Incidence/ or incidence.mp.	1114028
20	mask*.mp.	113702
21	exp Methodology/	5410519
22	exp Mortality/ or mortality.mp.	1424552
23	exp Observational Study/	232492
24	exp Placebo/ or placebos.mp.	315236
25	predict*.mp.	2222147
26	exp Prevalence/ or prevalence.mp.	1078206
27	prognos*.mp.	1080028
28	exp Randomized Controlled Trial/ or exp Single Blind Procedure/	625804
29	random sample/ or random.mp.	324115
30	research design.mp.	32845
31	risk.mp.	3847182
32	singl*.mp.	2168292
33	trebl*.mp.	421
34	tripl*.mp.	173021
35	volunteer*.mp.	223872
36	(health* adj2 survey*).mp.	240229
37	or/7-36	17490212
38	5 and 6 and 37	4461
39	38 not ((exp animal/ or nonhuman/) not exp human/)	3719
40	limit 39 to yr="2014 -Current"	1851

Search date: 19 May 2021

Table 47 Search strategy for Latin American and Caribbean Health Sciences Literature (LILACS) 1982 to 24 May 2021

Line number	Search terms	Number of results
1	"WATER" AND (FLUORIDE or FLUORINE) or (FLURIDE OR FLURINE) OR ("FLUORINATION" OR FLUORIDATION) [Subject descriptor] and "WATER" AND (FLUORIDE or FLUORINE) or (FLURIDE OR FLURINE) OR ("FLUORINATION" OR FLUORIDATION) [Words] and ("CLINICAL TRIAL" or "CLINICAL TRIAL, PHASE I" or "CLINICAL TRIAL, PHASE II" or "CLINICAL TRIAL, PHASE III" or "CLINICAL TRIAL, PHASE IV" or "COMPARATIVE STUDY" or "CONSENSUS DEVELOPMENT CONFERENCE" or "CORRECTED AND REPUBLISHED ARTICLE" or "RANDOMIZED CONTROLLED TRIAL") or "META-ANALYSIS" [Publication type]	253

Search date: 19 May 2021

Table 48 Search strategy for Cochrane Library 1993 to 19 May 2021

Line number	Search terms	Number of results
1	MeSH descriptor: [Fluorides] explode all trees	2656
2	MeSH descriptor Fluoridation this term only	3
3	MeSH descriptor: [Fluorine] explode all trees	85
4	#1 or #2 or #3	2737
5	MeSH descriptor: [Water Supply] explode all trees	180
6	MeSH descriptor: [Water] explode all trees	2395
7	(water treatment in All Text or water fluorid* in All Text or community water in All Text)	1271
8	#5 or #6 or #7	3784
9	#4 and #8 with Cochrane Library publication date Between Jan 2014 and May 2021	23
10	#9 Apply Limits: Cochrane Reviews only	10

Search date: 19 May 2021

Table 49 Search strategy for Epistemonikos 1986 to 19 May 2021

Line number	Search terms	Number of results
1	(title:(fluoride AND water fluoridation AND review) OR abstract:(fluoride AND water AND fluoridation AND review))	21

Search date: 19 May 2021

Appendix E Supplementary grey literature search

General scoping searches were carried out in the search engine Google.com to gain an initial idea of terminology and likely key terms. Initial search terms used included combinations of water, fluoride, fluoridation, and community water fluoridation. Further searches were carried out using the websites of national and international dental and health organisations (see Table 50).

Table 50 Websites included in supplementary grey literature search

Organisation	Website
Australia	
National Health and Medical Research Council (NHMRC) AUS	https://www.nhmrc.gov.au/
Canada	
Canadian Dental Association	https://www.cda-adc.ca/en/index.asp
Canadian Institute for Health Information	https://www.cihi.ca/en
Ireland	
Dental Council (Ireland)	http://www.dentalcouncil.ie
Environmental Protection Agency	https://www.epa.ie/
Irish Expert Body on Fluorides and Health	https://www.fluoridesandhealth.ie/
New Zealand	
Ministry of Health (NZ)	https://www.health.govt.nz/
Environmental Health Intelligence New Zealand (EHINZ)	https://www.ehinz.ac.nz/
United Kingdom	
NICE	https://www.nice.org.uk/
Scottish Dental Clinical Effectiveness Programme	https://www.sdcep.org.uk
British Dental Association	https://www.bda.org/
United States of America	
Center for Disease Control (CDC)	https://www.cdc.gov/fluoridation/
U.S. Department of Health and Human Services	https://www.hhs.gov/about/index.html
American Dental Association (ADA)	https://www.ada.org/
International Bodies	
European Food Safety Authority	https://www.efsa.europa.eu/en
International HTA Database	https://database.inahta.org/
World Health Organization	https://www.who.int/

Appendix F Extraction form

Data on the following parameters were extracted:

- Study author(s)
- Year of publication
- Research question
- Primary study design, designated by review team according to definitions by Hennekens and Buring [39] (see Glossary of terms); Randomised controlled trial/controlled clinical trial/retrospective/prospective cohort study/case-control study/cross-sectional survey/ecological/correlational study
- Study country
- Length of study period
- Study exposure(s) for observational studies, or cases for case-control studies
- Length of exposure
- Study comparator(s)
- Study outcome(s)
- Sample size recruited
- Sample size for analysis
- Mean age in years
- Gender (% female)
- Detailed results, including units and method of measurement, number of participants for analysis, statistical method, summary and variability statistics, confounding variables controlled for within design and analysis, significance, and narrative summary of findings

Appendix G Quality assessment tools

Ecological, cross-sectional, and cohort studies

Ecological, cross-sectional, and cohort studies were assessed for methodological quality by a single reviewer using the National Heart, Lung, and Blood Institute's (NHLBI's) quality assessment tool for observational cohort and cross-sectional studies [40]. The 14 items in this tool are displayed in Table 51. For each study, an overall quality rating was calculated using a bespoke system, based on essential criteria for high-quality observational and cohort studies [39]. Five items (bolded in Table 51 below) were selected and scored as outlined below.

Table 51 National Heart, Lung, and Blood Institute quality assessment tool for observational cohort and cross-sectional studies

Item	Scoring
1. Was the research question or objective in this paper clearly stated?	-
2. Was the study population clearly specified and defined?	-
3. Was the participation rate of eligible persons at least 50%?	**Yes 1.0** **No 0.0**
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	**Yes 1.0** **No 0.0**
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	-
5A. Was a sample size justification, power description, or variance and effect estimates provided?	**Yes 1.0** **Partly 0.5** **No 0.0**
5B. Was a description of variance provided?	-
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	-
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	-
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	-
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	-
10. Was the exposure(s) assessed more than once over time?	-
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	-
12. Were the outcome assessors blinded to the exposure status of participants?	-
13. Was loss to follow-up after baseline 20% or less?	**Yes 1.0** **No 0.0**
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	**Almost all 1.0** **Partial 0.5** **Limited 0.0** **None 0.0**
Quality rating	

Note: Responses of "Not reported" and "Not applicable" were scored 0.0 for each item.

For item 14, key potential confounding variables were identified based on established risk factors for the condition or outcome under consideration (see Appendix A); while some studies controlled for a large number of variables in their models, only these key confounding variables were considered for item 14. The key confounding variables for each condition or outcome of interest are displayed in Table 52.

Table 52 Key confounding variables

Condition/outcome	Key confounding variables
Bone mineral density/fractures	Sex, age, family history of osteoporosis, heavy alcohol and tobacco use
IQ	Parental socioeconomic status, maternal education, nutrition, perinatal environment, genetic factors, exposure to lead or arsenic, low birth weight
ADHD	Brain injury, exposure to neurotoxic chemicals (e.g. lead, manganese, arsenic, mercury, poly-chlorinated biphenyls, food additives and colourings), alcohol and tobacco use during pregnancy, premature delivery, low birth weight
Osteosarcoma/Ewing sarcoma	Previous radiation therapy, other bone disorders, viral diseases
Bladder cancer	Sex, age, alcohol and tobacco use, diet and lifestyle factors, catheter use
Secondary and all cancers	We could find no evidence or rationale for a link between fluoride exposure and secondary or all cancers; therefore, beyond the basic confounding factors of sex, age, alcohol and tobacco use, and diet and lifestyle factors, we cannot determine whether confounding factors have been wholly or partially accounted for.
Hypothyroidism and levels of thyroid-stimulating hormones	Sex, age, family history, iodine deficiency, radiation exposure or therapy, some medications
Sleep disturbance due to reduced functioning of pineal gland	We could find no strong rationale for a link between fluoride exposure and sleep disturbances related to pineal gland functioning; therefore, we cannot determine whether confounding factors have been wholly or partially accounted for.
Renal conditions	Sex, race, obesity, nutrition, inflammatory bowel diseases, some medications
Birth or birthing abnormalities	Maternal age, genetic factors, pregnancy and labour complications, maternal smoking, maternal alcohol and drug use, maternal medical history (including diabetes), some infections (e.g. toxoplasmosis), exposure to lead, pesticides, or radiation
Infant abnormalities	Sleeping position, use of soft or loose bedding or unsafe beds, maternal smoking during pregnancy, premature birth or low birth weight, exposure to second-hand smoke
All-cause mortality	We could find no evidence or rationale for a link between fluoride exposure and all causes of mortality; therefore, we cannot determine whether confounding factors have been wholly or partially accounted for.

For each study, the scores were summed (for a total score ranging from 0.0 to 5.0). Studies scoring less than 3.0 were rated 'low quality', studies scoring 3.0 were rated 'moderate quality', and studies scoring 3.5 or more were rated 'high quality'. As many studies were cross-sectional in nature (point-in-time surveys) and scored 0.0 on item 13 (loss to follow-up not applicable), the maximum possible score for these studies was effectively capped at 4.0; for this reason, the threshold for 'high quality' was set at 3.5, rather than 4.0, in order to allow more effective differentiation of studies at the upper end of the range of scores.

Case-control studies

Case-control studies were assessed for methodological quality by a single reviewer using the NHLBI's quality assessment tool for case-control studies.[40] The 13 items in this tool are displayed in Table 53. For each study, an overall quality rating was calculated using a bespoke system, based on essential criteria for high-quality case-control studies [39]. Five items (bolded in Table 53 below) were selected and scored as outlined below. Item 13 was scored according to the same key confounding variables as outlined for item 14 for the ecological, cross-sectional, and cohort studies. The same thresholds for low (<3.0), moderate (3.0) and high (3.5) quality were set as for the ecological, cross-sectional, and cohort studies.

Table 53 National Heart, Lung, and Blood Institute quality assessment tool for case-control studies

Item	
1. Was the research question or objective in this paper clearly stated and appropriate?	-
2. Was the study population clearly specified and defined?	-
3. Was an appropriate target population clearly defined per the research question? Did the cases adequately represent the cases that arose in the target population?	-
4. Did the authors include a sample size justification?	**Yes 1.0** **No 0.0**
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	**Yes 1.0** **Partly 0.5** **No 0.0**
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	**Yes 1.0** **No 0.0**
7. Were the cases clearly defined and differentiated from controls?	-
8. If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?	-
9. Was there use of concurrent controls?	-
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	**Yes 1.0** **Partly 0.5** **No 0.0**
11. Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (including the same time period) across all study participants?	-
12. Were the assessors of exposure/risk blinded to the case or control status of participants?	-
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	**Almost all 1.0** **Partial 0.0** **Limited 0.0** **None 0.0**
Quality rating	

Note: Responses of “Not reported” and “Not applicable” were scored 0.0 for each item.

Appendix H Feasibility assessment for meta-analysis

Bone health

Bone characteristics

Of 18 outcomes, 10 were examined in only one paper (total body, lumbar spine, lumbar spine L1-L4, estimated volumetric L3, anterior-posterior lumbar spine, total hip, femoral head, proximal femur, bone stiffness, and osteoporosis).

Four outcomes (radius distal, radius proximal, radius calcaneus, and intertrochanter) were examined by two cross-sectional papers (Cauley *et al.* (1995) and Phipps *et al.* (2000)). However, both papers are based on data from the Study of Osteoporotic Fractures Pittsburgh. The later paper presents an updated phase of the study, carried out with a larger sample over a longer period of time, with the intention of increasing power. For this reason, as some participants are likely to be counted in both analyses, it was determined that these papers should not be eligible for inclusion in the same meta-analysis. Only the analysis by Phipps *et al.* was included in this feasibility analysis, as the methods and demographics were comparable, but the sample size was larger than in the analysis by Cauley *et al.*

For the remaining four outcomes, a feasibility assessment for meta-analysis was completed.

Lumbar spine L2-L4

Two studies examined BMD of the lumbar spine L2-L4 (Kröger *et al.* (1994) and Lehmann *et al.* (1998)).

1. The studies were compatible on study design; both were cross-sectional studies.
2. The studies were compatible on outcome measurement, with both assessing BMD using dual-energy X-ray absorptiometry.
3. The studies were not compatible on population. Although both provided data specifically for female populations, the age profiles of the participants were substantially different (see Table 54), with the ages of the participants in the earlier study being older and less variable. For this reason, it was determined that meta-analysis of the findings for these two studies would be inappropriate.

Table 54 Age profiles for participants in studies of BMD

	Fluoridated group Mean age (SD)	Non-fluoridated group Mean age (SD)
Kröger *et al.* (1994)	53.2 (2.8)	53.5 (2.8)
Lehmann *et al.* (1998)	42.9 (12.1)	37.8 (12.8)
Phipps *et al.* (2000)	73.9	74.5

Femoral neck

Four papers examined BMD of the femoral neck (Kröger *et al.* (1994), Cauley *et al.* (1995), Lehmann *et al.* (1998), Phipps *et al.* (2000)).

As outlined above, the papers by Cauley *et al.* and Phipps *et al.* were not eligible for inclusion in the same meta-analysis. Only the analysis by Phipps *et al.* was included in this feasibility analysis, as the methods and demographics were comparable, but the sample size was larger than in the analysis by Cauley *et al.*

1. The remaining three studies were compatible on study design; all were cross-sectional studies.

2. The three studies were compatible on outcome assessment, with Kröger *et al.* and Lehmann *et al.* using dual-energy X-ray absorptiometry, and Phipps *et al.* using single photon absorptiometry, which yields comparable results.
3. The studies were not compatible on population. Although all provided data specifically for female populations, the age profiles of the participants were substantially different (see Table 54). The age profile of the participants in the Kröger *et al.* study was older and less variable than that of participants in the Lehmann study. Participants in the Phipps *et al.* study were notably older again. For this reason, it was determined that meta-analysis of the findings for these two studies would be inappropriate.

Ward's triangle

Three papers examined BMD of Ward's triangle (Cauley *et al.* (1995), Lehmann *et al.* (1998), Phipps *et al.* (2000)).

As outlined above, the papers by Cauley *et al.* and Phipps *et al.* were not eligible for inclusion in the same meta-analysis. Only the analysis by Phipps *et al.* was included in this feasibility analysis, as the methods and demographics were comparable but the sample size was larger than in the analysis by Cauley *et al.*

1. The remaining two studies were compatible on study design; both were cross-sectional studies.
2. The two studies were not compatible on outcome assessment; while Lehmann *et al.* used using dual-energy X-ray absorptiometry, the method used by Phipps *et al.* was not specified for this outcome.
3. Furthermore, the studies were not compatible on population. Although both provided data specifically for female populations, the age profiles of the participants in the Phipps *et al.* study were substantially older (see Table 54). For this reason, it was determined that meta-analysis of the findings for these two studies would be inappropriate.

Trochanter

Three papers examined BMD of the trochanter (Cauley *et al.* (1995), Lehmann *et al.* (1998), Phipps *et al.* (2000)).

As outlined above, the papers by Cauley *et al.* and Phipps *et al.* were not eligible for inclusion in the same meta-analysis. Only the analysis by Phipps *et al.* was included in this feasibility analysis, as the methods and demographics were comparable but the sample size was larger than in the analysis by Cauley *et al.*

1. The remaining two studies were compatible on study design; both were cross-sectional studies.
2. The two studies were not compatible on outcome assessment; while Lehmann *et al.* used dual-energy X-ray absorptiometry, the method used by Phipps *et al.* was not specified for this outcome.
3. Furthermore, the studies were not compatible on population. Although both provided data specifically for female populations, the age profile of the participants in the Phipps *et al.* study was substantially older (see Table 54). For this reason, it was determined that meta-analysis of the findings for these two studies would be inappropriate.

Fractures

Of nine fracture types, five were examined in only one paper (all fractures, osteoporotic fractures, ankle fractures, non-wrist/non-ankle fractures, humerus fractures).

Two fracture types (incidental spinal/vertebral fractures and non-spine/non-vertebral fractures) were examined by two cross-sectional papers (Cauley *et al.* (1995) and Phipps *et al.* (2000)). However, as outlined above, the papers by Cauley *et al.* and Phipps *et al.* were not eligible for inclusion in the same meta-analysis. Therefore, no meta-analysis was possible for studies of these fracture types.

For the two remaining outcomes, a feasibility assessment for meta-analysis was completed.

Hip fracture

Nine papers examined incidence of hip fracture. Seven were ecological studies and therefore not suitable for inclusion in meta-analysis.

The remaining two papers were based on one cross-sectional study (Cauley *et al.* (1995) and Phipps *et al.* (2000)). As outlined above, these two papers were not eligible for inclusion in the same meta-analysis. Therefore, no meta-analysis was possible for studies of hip fracture.

Wrist fracture

Three papers examined incidence of wrist fracture (Kröger *et al.* (1994), Cauley *et al.* (1995), and Phipps *et al.* (2000)).

As outlined above, the papers by Cauley *et al.* and Phipps *et al.* were not eligible for inclusion in the same meta-analysis. Only the analysis by Phipps *et al.* was included in this feasibility analysis, as the methods and demographics were comparable but the sample size was larger than in the analysis by Cauley *et al.*

1. The remaining two studies were compatible on study design; both were cross-sectional studies.
2. The two studies were compatible on outcome assessment, both using self-reported fractures, although Phipps *et al.* excluded fractures due to major trauma (e.g. motor vehicle accident).
3. The studies were not compatible on population. Although both provided data specifically for female populations, the age profiles of the participants in the Phipps study were substantially older (see Table 54). For this reason, it was determined that meta-analysis of the findings for these two studies would be inappropriate.

Neuropsychological outcomes

IQ/neuropsychological development

Five papers examined outcomes related to IQ/neuropsychological development. Three of these papers (Green *et al.* (2019), Till *et al.* (2020), and Farmus *et al.* (2021)) analysed data from the same programme of research (601 mother-child pairs enrolled in the Maternal-Infant Research on Environmental Chemicals longitudinal cohort study). For this reason, it was determined that these papers should not be eligible for inclusion in the same meta-analysis. The feasibility assessment therefore assessed the appropriateness of synthesising the findings from the studies by Broadbent *et al.* (2015) and Ibarluzea *et al.* (2021) with any of these three analyses of the cohort study.

1. The studies were compatible on study design; all were prospective cohort studies. However, the studies by Green *et al.*, Till *et al.*, Farmus *et al.*, and Ibarluzea *et al.* assigned fluoridation status on an ecological/population basis.
2. The studies were not compatible on outcome assessment; Broadbent *et al.* used the Weschler Adult Intelligence Scale Revised and its subscales, Ibarluzea *et al.* used the Bayley Mental Development Index and McCarthy Scales of Children's Abilities, and Green *et al.*, Till *et al.* and

Farmus *et al.* all used the Weschler Pre-school and Primary Scale of Intelligence and its subscales.

3. Furthermore, the studies were not compatible on population, as the age profiles of participants were substantially different. Broadbent *et al.* assessed IQ at age 7-13 and age 38, while Ibarluzea *et al.* assessed at age 4, and Green *et al.*, Till *et al.*, and Farmus *et al.* assessed at age 3-4.

For these reasons, it was determined that meta-analysis of the findings for these studies would be inappropriate.

Neurodevelopmental disorders

Both papers on neurodevelopmental disorders (Barberio *et al.* (2017a) and Riddell *et al.* (2019)) analysed data from the same dataset, Cycles 2 (2009–2011) and 3 (2012–2013) of Statistics Canada's Canadian Health Measures Survey. For this reason, it was determined that meta-analysis of the findings for these two papers would be inappropriate.

Cancer

Bone cancer

Eleven papers examined incidence of bone cancers. Eight were ecological studies and therefore not suitable for inclusion in meta-analysis.

Of the remaining three papers, the 2020 paper by Kim *et al.* and the 2006 paper by Bassin *et al.* both present data from the Harvard Fluoride Osteosarcoma Study. Kim *et al.* (2006) incorporates a broader population and other variables of interest. For this reason, as some participants are likely to be counted in both analyses, it was determined that these analyses should not be eligible for inclusion in the same meta-analysis. Only the paper by Kim *et al.* was included in this feasibility analysis, as the methods and demographics were comparable but the sample size was larger than in the original paper by Bassin *et al.*

1. The remaining two studies (McGuire *et al.* (1991) and Kim *et al.* (2020)) were compatible on study design; both were case-control studies.
2. The studies were compatible on outcome assessment; both examined cases of osteosarcoma within given treating departments.
3. The studies were compatible on population; although McGuire *et al.* recruited only patients diagnosed before the age of 40, the distribution of age at diagnosis was similar for the cases in each study, as was the gender ratio. However, the studies differed in the method of matching; although both studies matched cases and controls based on age and gender, McGuire *et al.* matched location based on county of residence, while Kim matched location based on distance from the hospital. In addition, only McGuire *et al.* used a 1:1 ratio of matching.
4. The studies were compatible on both exposure and comparator; both studies collected residential histories and estimated level of fluoride in the drinking water at each address of residence for both cases and controls.

Due to the different approaches to matching in the studies, it was determined that meta-analysis of the findings for these studies would be inappropriate.

Other cancers

Three other cancer outcomes (incidence of all cancers, bladder cancer, and secondary bone cancer) were each examined in only one study.

Endocrine conditions

Three papers examined outcomes related to thyroid functioning and one study examined outcomes related to sleep disturbance. A majority of the specific outcomes (hyperthyroidism prevalence, diagnosed thyroid condition, less than/more than recommended sleep duration, trouble sleeping, daytime sleepiness) were examined by only one study.

TSH levels were examined by two papers (Barberio *et al.* (2017b) and Malin *et al.* (2018)). However, both papers analyse data from the same dataset, Cycles 2 (2009–2011) and 3 (2012–2013) of Statistics Canada's Canadian Health Measures Survey. For this reason, it was determined that meta-analysis of the findings for these two analyses would be inappropriate.

Renal conditions

Renal conditions (renal calculi) were examined in only one ecological study.

Birth or birthing abnormalities

All three studies of birth or birthing abnormalities were ecological studies and therefore not suitable for inclusion in meta-analysis.

Infant abnormalities

Infant abnormalities (SIDS) were examined in only one case-control study.

All-cause mortality

All-cause mortality was examined in only one ecological study.

Appendix I Overview of literature search results

Table 55 Overview of literature search results

Databases	Date	Number of results
Ovid MEDLINE	19 May 2021	2154
Ovid Embase	19 May 2021	1838
Cochrane Library (John Wiley & Sons Inc)	19 May 2021	13
Latin American and Caribbean Health Sciences Literature (LILACS)	24 May 2021	223
Epistemonikos (Epistemonikos Foundation)	19 May 2021	21
PROSPERO (University of York)	19 May 2021	20
Cochrane CENTRAL (John Wiley & Sons Inc)	19 May 2021	10
Total before deduplication		**4279**
Total after deduplication		**3259**
Total retained for analysis after screening		**37**
Total added from reference chasing		**1**

Appendix J Included studies

Arnold CM, Bailey DA, Faulkner RA, McKay HA, McCulloch RG. The effect of water fluoridation on the bone mineral density of young women. Canadian journal of public health. 1997 Nov;88(6):388-91.

Barberio AM, Hosein FS, Quiñonez C, McLaren L. Fluoride exposure and indicators of thyroid functioning in the Canadian population: implications for community water fluoridation. J Epidemiol Community Health. 2017 Oct 1;71(10):1019-25.

Barberio AM, Quinonez C, Hosein FS, McLaren L. Fluoride exposure and reported learning disability diagnosis among Canadian children: Implications for community water fluoridation. Canadian Journal of Public Health. 2017 May;108(3):e229-39.

Bassin EB, Wypij D, Davis RB, Mittleman MA. Age-specific fluoride exposure in drinking water and osteosarcoma (United States). Cancer Causes & Control. 2006 May;17(4):421-8.

Blakey K, Feltbower RG, Parslow RC, James PW, Gómez Pozo B, Stiller C, Vincent TJ, Norman P, McKinney PA, Murphy MF, Craft AW. Is fluoride a risk factor for bone cancer? Small area analysis of osteosarcoma and Ewing sarcoma diagnosed among 0–49-year-olds in Great Britain, 1980–2005. International Journal of Epidemiology. 2014 Feb 1;43(1):224-34.

Broadbent JM, Thomson WM, Ramrakha S, Moffitt TE, Zeng J, Foster Page LA, Poulton R. Community water fluoridation and intelligence: prospective study in New Zealand. American journal of public health. 2015 Jan;105(1):72-6.

Cauley JA, Buhari AM, Murphy PA, Riley TJ. Effects of fluoridated drinking water on bone mass and fractures: the study of osteoporotic fractures. Journal of bone and mineral research. 1995 Jul;10(7):1076-86.

Chachra D, Limeback H, Willett TL, Grynpas MD. The long-term effects of water fluoridation on the human skeleton. Journal of dental research. 2010 Nov;89(11):1219-23.

Comber H, Deady S, Montgomery E, Gavin A. Drinking water fluoridation and osteosarcoma incidence on the island of Ireland. Cancer Causes & Control. 2011 Jun;22(6):919-24.

Crnosija N, Choi M, Meliker JR. Fluoridation and county-level secondary bone cancer among cancer patients 18 years or older in New York State. Environmental geochemistry and health. 2019 Apr;41(2):761-8.

Cunningham JE, McCague H, Malin AJ, Flora D, Till C. Fluoride exposure and duration and quality of sleep in a Canadian population-based sample. Environmental Health. 2021 Dec;20(1):1-0.

Danielson C, Lyon JL, Egger M, Goodenough GK. Hip fractures and fluoridation in Utah's elderly population. Jama. 1992 Aug 12;268(6):746-8.

Dick AE, Ford RP, Schluter PJ, Mitchell EA, Taylor BJ, Williams SM, Stewart AW, Becroft DM, Thompson JM, Scragg R, Hassall IB. Water fluoridation and the sudden infant death syndrome. The New Zealand medical journal. 1999 Aug 1;112(1093):286-9.

Farmus L, Till C, Green R, Hornung R, Martinez-Mier EA, Ayotte P, Muckle G, Lanphear B, Flora D. Critical windows of fluoride neurotoxicity in Canadian children. Environmental Research. 2021 May 27:111315.

Green R, Lanphear B, Hornung R, Flora D, Martinez-Mier EA, Neufeld R, Ayotte P, Muckle G, Till C. Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada. JAMA Pediatrics. 2019 Oct 1;173(10):940-8.

Hrudey SE, Soskolne CL, Berkel JO, Fincham SH. Drinking water fluoridation and osteosarcoma. Canadian Journal of Public Health. 1990;81(6):415-6.

Ibarluzea J, Gallastegi M, Santa-Marina L, Zabala AJ, Arranz E, Molinuevo A, *et al.* Prenatal exposure to fluoride and neuropsychological development in early childhood: 1-to 4 years old children. Environmental Research. 2021;112181.

Jacobsen SJ, Goldberg J, Cooper C, Lockwood SA. The association between water fluoridation and hip fracture among white women and men aged 65 years and older: a national ecologic study. Annals of epidemiology. 1992 Sep 1;2(5):617-26.

Jacobsen SJ, O'Fallon WM, Melton 3rd LJ. Hip fracture incidence before and after the fluoridation of the public water supply, Rochester, Minnesota. American journal of public health. 1993 May;83(5):743-5.

Kim FM, Hayes C, Burgard SL, Kim HD, Hoover RN, National Osteosarcoma Etiology Group, Douglass CW, Couper D. A case-control study of fluoridation and osteosarcoma. Journal of dental research. 2020 Sep;99(10):1157-64.

Kröger H, Alhava E, Honkanen R, Tuppurainen M, Saarikoski S. The effect of fluoridated drinking water on axial bone mineral density—a population-based study. Bone and mineral. 1994 Jan 1;27(1):33-41.

Lee N, Kang S, Lee W, Hwang SS. The association between community water fluoridation and bone diseases: A natural experiment in Cheongju, Korea. International Journal of Environmental Research and Public Health. 2020 Jan;17(24):9170.

Lehmann R, Wapniarz M, Hofmann B, Pieper B, Haubitz I, Allolio B. Drinking water fluoridation: Bone mineral density and hip fracture incidence. Bone. 1998 Mar 1;22(3):273-8.

Lowry R, Steen N, Rankin J. Water fluoridation, stillbirths, and congenital abnormalities. Journal of Epidemiology & Community Health. 2003 Jul 1;57(7):499-500.

Mahoney MC, Nasca PC, Burnett WS, Melius JM. Bone cancer incidence rates in New York State: time trends and fluoridated drinking water. American Journal of Public Health. 1991 Apr;81(4):475-9.

Malin AJ, Riddell J, McCague H, Till C. Fluoride exposure and thyroid function among adults living in Canada: effect modification by iodine status. Environment international. 2018 Dec 1;121:667-74.

McGuire SM, Vanable ED, McGuire MH, Buckwalter JA, Douglass CW. Is there a link between fluoridated water and osteosarcoma?. The Journal of the American Dental Association. 1991 Apr 1;122(4):38-45.

National Fluoridation Information Service (NFIS), 2013. Community Water Fluoridation Community Water Fluoridation and Osteosarcoma - Evidence from Cancer Registries. Available at: http://www.rph.org.nz/content/45a27238-502a-48ed-a90c-aa306b3ac449.cmr .

O' Sullivan V, O' Connell BC. Water fluoridation, dentition status and bone health of older people in Ireland. Community dentistry and oral epidemiology. 2015 Feb;43(1):58-67.

Cohn PD. An epidemiologic report on drinking water and fluoridation. New Jersey, Department of Health, Environmental Health Service; 1992.

Peckham S, Lowery D, Spencer S. Are fluoride levels in drinking water associated with hypothyroidism prevalence in England? A large observational study of GP practice data and fluoride levels in drinking water. J Epidemiol Community Health. 2015 Jul 1;69(7):619-24.

Phipps KR, Orwoll ES, Mason JD, Cauley JA. Community water fluoridation, bone mineral density, and fractures: prospective study of effects in older women. Bmj. 2000 Oct 7;321(7265):860-4.

Riddell JK, Malin AJ, Flora D, McCague H, Till C. Association of water fluoride and urinary fluoride concentrations with attention deficit hyperactivity disorder in Canadian youth. Environment international. 2019 Dec 1;133:105190.

Suarez-Almazor ME, Flowerdew G, Saunders LD, Soskolne CL, Russell AS. The fluoridation of drinking water and hip fracture hospitalization rates in two Canadian communities. American journal of public health. 1993 May;83(5):689-93.

Till C, Green R, Flora D, Hornung R, Martinez-Mier EA, Blazer M, Farmus L, Ayotte P, Muckle G, Lanphear B. Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. Environment international. 2020 Jan 1;134:105315.

Young N, Newton J, Morris J, Morris J, Langford J, Iloya J, Edwards D, Makhani S, Verne J. Community water fluoridation and health outcomes in England: a cross-sectional study. Community dentistry and oral epidemiology. 2015 Dec;43(6):550-9.

Zhang X, Lu E, Stone SL, Diop H. Dental cleaning, community water fluoridation and preterm birth, Massachusetts: 2009–2016. Maternal and child health journal. 2019 Apr;23(4):451-8.

Appendix K Community water fluoridation schemes by country

Australia

Description of scheme: CWF in Australia began in the 1960s and 1970s following the acceptance of its benefits in December 1953. By 1984 almost 66% of the Australian population represented by 850 towns and cities had access to fluoridated drinking water. As of February 2017, Australia provides fluoridated water for 89% of the population in all states and territories have fluoridated water.

The first town to fluoridate the water supply in Australia was Beaconsfield, Tasmania in 1953. Queensland became the last state to formally require the addition of fluoride to public drinking water supplies in December 2008. Some areas within Australia have natural fluoride levels in the groundwater, which was estimated in 1991 to provide drinking water to approximately 0.9% of the population.

Quantity: 0.7 to 0.9 milligrams per litre 1.0 milligram per litre established in the 1966 Legislation. Optimum level for the Perth metropolitan area is 0.9 milligram per litre, with a range of 0.7 to 1.0 milligrams per litre.

Sources:

1. National Health and Medical Research Council. Water Fluoridation and Human Health in Australia: Questions and Answers [Internet]. Australian Government, Department of Health and Ageing. Available from: https://www.nhmrc.gov.au/sites/default/files/documents/attachments/water-fluoridationqa.pdf

Canada

Description of scheme: In Canada, CWF is the responsibility of the local government but is regulated by provincial, territorial, and federal governments. As a result, local governments across Canada have the option of fluoridating their water. In 1945, Brantford, Ontario, became the first Canadian city to fluoridate its water supplies. Toronto approved water fluoridation in 1955, but implementation was delayed until 1963 due to a campaign against it. The city continues to fluoridate its water today.

Fluoridation rates are highest in Ontario, Alberta, and Manitoba, at 70–75%. Quebec (about 6%), British Columbia (3.7%), and Newfoundland and Labrador (1.5%) have the lowest fluoridation rates, while Yukon and Nunavut have no fluoridation at all.

In 2007, approximately 45% of the Canadian population had access to fluoridated water supplies. The Region of Waterloo held a non-binding referendum in 2010 to determine whether water fluoridation should continue. The vote resulted in 50% voting against fluoridation. The vote was honoured by the regional council, and over forty years of fluoridation in the city of Waterloo (Ontario) came to an end in November 2010.

The decision to stop water fluoridation has also been replicated in other cities. In 2011, Calgary city council (Alberta) voted 10–3 to stop adding fluoride to the city's drinking water, having started water fluoridation in 1991.

Lakeshore and Amherstberg (both in Ontario) have voted to stop fluoridating their water. In Ontario, the cities of Hamilton, London, and Toronto have chosen to continue fluoridation.

On January 28, 2013, Windsor city council (in Ontario) voted 8–3 to discontinue fluoridation of Windsor's drinking water for five years, honouring a February 2012 recommendation from the Windsor Utilities Commission. On 14 December 2018, Windsor city council voted 8–3 on December 14, 2018 to reintroduce fluoridation of the city's drinking water. According to the health unit's Oral Health 2018 report, the percentage of children with tooth decay or requiring urgent care increased by 51% in 2016–17 compared with 2011–12.

Quantity: In 2008 the recommended fluoride levels in Canada were reduced from 0.8 to 1.0 mg/L to 0.7 mg/L to minimise the risk of dental fluorosis.

Sources:

1. Rabb-Waytowich D. Water fluoridation in Canada: past and present. J Can Dent Assoc. 2009;75(6).

England

Description of scheme: The first and largest UK water fluoridation scheme was introduced in Birmingham in 1964. Further territories in the Midlands and the North of England were added during the next two decades.

Fluoridated water is now provided to about 5.8 million people in England, with the majority (92%) receiving it as part of a fluoridation project. The remaining 8% get it because it's naturally found in some water sources. The decision about whether to add fluoride to the water supply is made by individual local authorities. Water fluoridation schemes are currently in place in the West Midlands, the Northeast, the East Midlands, Eastern England, the Northwest, Yorkshire, and the Humber. In some parts of the country, such as the Northeast and the Midlands, the public water supply naturally contains fluoride levels comparable to those found in schemes. Some private water supplies have a higher concentration

Quantity: Around 1 mg of fluoride per litre of water

Sources:

1. Fluoride [Internet]. nhs.uk. 2021. Available from: https://www.nhs.uk/conditions/fluoride/
1. Lowery G, Bunn S. Water fluoridation and dental health [Internet]. POST. 2021. Available from: https://post.parliament.uk/water-fluoridation-and-dental-health/

Finland

Description of scheme: Kuopio is the only city in Finland with a population of at least 70,000 people that has ever received fluoridated water. Between 1959 and 1992, the drinking water in the city of Kuopio was fluoridated up to a concentration of 1.2 mg/l to prevent dental caries. The natural fluoride content of this city's drinking water is low (0-0.3 mg/1).

Quantity: 1.2 mg/l

Sources:

1. Kröger H, Alhava E, Honkanen R, Tuppurainen M, Saarikoski S. The effect of fluoridated drinking water on axial bone mineral density—a population-based study. Bone and mineral. 1994 Jan 1;27(1):33-41.

Germany

Description of scheme: CWF schemes were introduced in Germany in 1961 in Germany and were phased out in West Germany in 1971 and East Germany in 1999. In East Germany, drinking water fluoridation was restricted to a few communities.

Quantity: 1 mg/l

Sources:

1. Lehmann R, Wapniarz M, Hofmann B, Pieper B, Haubitz I, Allolio B. 3rinking water fluoridation: Bone mineral density and hip fracture incidence. Bone. 1998 Mar 1;22(3):273-8.

Ireland

Description of scheme: One of the Department of Health's main public health interventions to prevent dental caries is the fluoridation of public piped water supplies in the Republic of Ireland currently at levels of 0.6 to 0.8 ppm. CWF at a level of 1 ppm began in Ireland in 1964 as a measure to prevent dental caries. A major review of Ireland's water fluoridation policy in 2002 showed an increasing occurrence of dental fluorosis. As a result, in 2007, the fluoride level in drinking water in Ireland was lowered to a range of 0.6 to 0.8 ppm, with a target of 0.7 ppm. In Ireland, statutory regulations for Fluoridation of Water Supplies stipulate that fluoride may be added to public water supplies, typically in the form of hydrofluorosilicic acid.

Quantity: 0.6 to 0.8 ppm since 2002

Sources:

1. O' Sullivan V, O' Connell BC. Water fluoridation, dentition status and bone health of older people in Ireland. Community dentistry and oral epidemiology. 2015 Feb;43(1):58-67.
2. Sutton M, Kiersey R, Farragher L, Long J. Health effects of water fluoridation. 2015. Available from: https://www.hrb.ie/fileadmin/publications_files/Health_Effects_of_Water_Fluoridation.pdf

New Zealand

Description of scheme: The use of water fluoridation first began in Hastings, New Zealand in 1954. In 1957, a Commission of Inquiry was convened, and by the mid-1960s, the use of CWF had rapidly expanded. New Zealand has fluoridated water supplied to about half of the total population. Of the six main centres, only Christchurch and Tauranga do not have a fluoridated water supply. The majority of Wellington's water supply is fluoridated, but the suburbs of Petone and Korokoro receive non-fluoridated water.

The Auckland suburbs of Onehunga and Huia Village do not fluoridate their water. A Hamilton City Council committee voted in 2013 to remove fluoride beginning in late June of that year. A referendum was held during the council elections in October 2013 with approximately 70% of voters voting for fluoride to be reintroduced into the water supply, and in March 2014, the council voted 9 to 1 to re-introduce fluoride into the supply. In a 2007 referendum, approximately half of voters in Central Otago, South Otago, and the Southland Region opposed fluoridation, and voters in the Waitaki District were against water fluoridation for all Wards. Ashburton and Greymouth also voted against fluoridation. In June 2018, the Supreme Court of New Zealand in New Health New Zealand Inc versus South Taranaki District Council upheld the legality of water fluoridation in New Zealand.

In November 2021, the Health (Fluoridation of Drinking Water) Amendment Bill was passed, shifting decision-making authority on water fluoridation from local authorities to the Director-General of Health.

Quantity: Optimal level of between 0.7ppm to 1ppm

Sources:

1. Community Water Fluoridation | New Zealand Dental Association [Internet]. Nzda.org.nz. Available from: https://www.nzda.org.nz/public/our-initiatives/community-water-fluoridation
2. Fluoridation: an update on Evidence [internet]. Pmcsa.ac.nz. Available from: https://www.pmcsa.ac.nz/topics/fluoridation-an-update-on-evidence/ (accessed 11 Feb 2022)
3. New Zealand Parliament. Health (Fluoridation of Drinking Water) Amendment Bill, 2021. 2021. Available from: https://www.parliament.nz/en/pb/bills-and-laws/bills-proposed-laws/document/00DBHOH_BILL71741_1/health-fluoridation-of-drinking-water-amendment-bill (accessed 11 Feb 2022)

South Korea

Description of scheme: In Korea, the first CWF programme was initiated in 1981 at Jinhae City.In 2002, It was expanded to include 32 local areas and 36 water treatment plants covering 8.9% of the Korean population. The oral health of this population has greatly improved as a result of this program, with the mean decayed, missing, and filled teeth index among Korean children steadily decreasing since the 1980s. By 2017, the CWF program's implementation areas had been reduced to 20 local areas and 24 water treatment plants, covering only 6.7 percent of the Korean population. Legislation requiring mandatory fluoridation was introduced in 2002, but its implementation has been postponed since then due to opposition from water companies, municipalities, and the public.

Quantity: Circa 0.8 ppm

Sources:

1. Kim HN, Kim JH, Kim SY, Kim JB. Associations of community water fluoridation with caries prevalence and oral health inequality in children. International journal of environmental research and public health. 2017 Jun;14(6):631.

Spain

Description of scheme: Over 4.25 million people (11% of the population) in Spain have either artificial or natural optimal community water fluoridation. The vast majority have access to artificial community water fluoridation. Areas of Spain with water fluoridation schemes are the Basque Country in the north (including the cities of Bilbao and San Sebastian); Girona in the north-east of Catalonia; parts of Murcia province in the south-east of Spain; and parts of Andalucía province (including the cities of Seville and Cordoba).

Quantity: Recommended 0.7 ppm

Sources:

1. British Fluoridation Society. The extent of water fluoridation [Internet]. 2012 [cited 2021 Sep 2]. (One in a Million: The Facts About Water Fluoridation). Available from: http://www.bfsweb.org/onemillion/09%20One%20in%20a%20Million%20-%20The%20Extent%20of%20Fluoridation.pdf

2. Vitoria I, Maraver F, Almerich-Silla JM. Flúor en aguas de consumo público españolas y prevención de la caries dental. Gac Sanit. 2014;28:255–6.

USA

Description of scheme: CWF and fluoride toothpaste are the most common sources of non-dietary fluoride in the USA. CWF began in 1945, reaching 49% of the population by 1975 and 67% by 2012. Fluoride toothpaste was first marketed in the United States in 1955. By the 1990s, fluoride toothpaste accounted for more than 90% of the toothpaste market. Fluoride is now available in mouth rinses, dietary fluoride supplements, and professionally applied fluoride compounds. In 2010, a study conducted by the Centers for Disease Control and Prevention in the United States found that 40.7 percent of adolescents aged 12–15 had dental fluorosis between 1999 and 2004. In response, the Department of Health and Human Services and the Environmental Protection Agency (EPA) proposed in 2011 to reduce the recommended level of fluoride in drinking water to the lowest end of the current range, 0.7 milligrams mg/L, from the previous recommended maximum of 0.7 to 1.2 mg/L, due to an increase in fluoride sources such as fluoridated toothpastes and mouthwashes.

Quantity: Decreased from 0.7 to 1.2 mg/L to 0.7 mg/L proposed in 2011 and implemented in 2014/5

Sources:

1. US Public Health Service. US Public health Service Recommendation for Fluoride Concentration in Drinking Water for the Preventin of Dental Caries. Reports and Recommendations. 2015;130(4):318-.

Appendix L Study characteristics

Table 56 Study characteristics for studies examining bone characteristics

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Bone mineral density									
Kröger *et al.* (1994)	Cross-sectional survey Finland	Kuopio Osteoporosis Study: all perimenopausal women aged 47–56 years residing in Kuopio Province, Finland in February 1989; random stratified sample of those willing to undergo bone densitometry	Investigated the influence of fluoridated drinking water on axial bone mineral density and the occurrence of fractures in a large perimenopausal population.	Women using fluoridated drinking water for more than 10 years (1.0–1.2 mg/L)	Women who did not have access to artificially fluoridated drinking water or who had used it for less than 10 years. The fluoride content of drinking water is low (0.0–0.3 mg/L).	BMD of spine and neck of femur	Fluoridated: 969; non-fluoridated: 2,253	Fluoridated 53.2; non-fluoridated 53.5	100
Cauley *et al.* (1995) (Pittsburgh) ●	Cross-sectional survey USA	Women aged 65 years or over, excluding black women (due to reduced incidence of hip fractures) and women unable to walk without the assistance of another person or who had	Evaluated whether women aged 65 years or over with greater lifetime exposure to residential fluoridated water have higher axial and	Years of exposure to fluoridated community water supplies recorded for each participant; exposure duration range: 1–38 years; mean fluoride	Zero years of exposure to fluoridated community water supplies recorded for each participant	Bone mineral content and density for the spine and hip and at the midpoint and ultradistal radius and calcaneus. Spinal and non-spinal	Zero years of fluoride exposure: 1,248; 1–10 years of fluoride exposure: 438; 11–20 years of fluoride exposure: 198; and	Zero years fluoride exposure 70.8, 1-10 years fluoride exposure 71.2, 11-20 years fluoride exposure 70.7, and >20 years fluoride exposure 71.6	100

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		bilateral hip replacements; most recruited from voter registration lists for ZIP codes within a 25-mile radius of Monessen, Pennsylvania, USA	appendicular bone mass and suffer fewer fractures than women compared with a lower lifetime exposure to residential fluoridated water.	concentration 1.01 ppm (±0.21 SD) for fluoridated public water		fractures were also recorded.	>20 years of fluoride exposure: 192		
Arnold *et al.* (1997)	Ecological or correlational study Canada	Females aged 18–25 years. All subjects had not travelled outside of their resident city in Canada for more than 4 years. Individuals with bone-affecting disorders, use of potential bone-affecting medications, long-term use of fluoride supplements, a history of amenorrhoea (fewer than three menses per year), and	Investigated the effect of long-term exposure to water fluoridation on bone mineral density in young adult women. The authors hypothesized that bone mineral density in the Saskatoon group would be higher than in the Regina group and that the greatest difference in bone mineral	Saskatoon, Saskatchewan, Canada, which has had supplemental fluoride in its water since 1954, at a level of approximately 1.0 mg/L; duration of exposure: >4 years	Regina, Saskatchewan, Canada, which has never had supplemental fluoride in its water supply and has a naturally occurring fluoride level of <0.12 to 0.15 mg/L in its water	BMD for the total body, lumbar spine, and proximal femur	Total: 57 (BMD fluoridated: 33; BMD non-fluoridated: 24)	Saskatoon (fluoridated) 21.3; Regina (non-fluoridated) 20.8	100

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		those who were currently pregnant were excluded.	density would be at the highly trabecular bone sites in the lumbar spine (anterior-posterior lumbar spine and volumetric Lumbar 3) as compared to the proximal femur.						
Lehmann *et al.* (1998)	Cross-sectional survey Germany	BMD analysis: volunteer employees at a local hospital in Germany, excluding those who used drugs affecting calcium metabolism and those with disease known to affect bone metabolism	Investigated the effect of drinking water fluoridation on bone mineral density by comparing two similar communities in the former German Democratic Republic, one with artificially fluoridated water at a level of 1 mg/L and one without.	Chemnitz, Germany, where water was fluoridated to 1 mg/L over a period of 30 years; duration of exposure: >10 years	Halle, Germany, where water was not fluoridated and had a naturally occurring fluoride level of 0.08–0.36 mg/L	BMD	555 (Chemnitz, Germany: 201 women and 41 men, totalling 242; Halle, Germany: 215 women and 98 men, totalling 313) [Authors Table 2]	Men in Chemnitz 42.9 (±12.1 SD); Halle 37.8 (±12.8 SD); Women in Chemnitz 40.7 (±11.3 SD); Halle 39.1 (±11.2 SD)	Chemnitz (fluoridated) 83.1; Halle (non-fluoridated) 68.7
Phipps *et al.* (2000) (Pittsburgh) ●	Cross-sectional survey USA	Prospective sample of 9,704 white women aged 65 years or	Determined whether older women with long term	Women exposed to fluoridated water	Women with no exposure to fluoridated water for the	BMD and fractures of the vertebrae,	No exposure: 3,218; continuous	No exposure 74.5, Continuous exposure 73.9	100

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		over in Portland, Oregon; Minneapolis, Minnesota; Baltimore, Maryland; and the Monongahela Valley, Pennsylvania, USA, recruited from jury selection and voter registration, motor vehicle records, and membership records of health plans. Excluded white women unable to walk without assistance and women who had bilateral hip replacement. Recruitment took place from 1986 to 1988.	exposure to fluoridated water had different bone mass and rates of fracture compared with women with no exposure.	continuously for the last 20 years; levels not specified, USA standard target fluoride level was 0.7–1.2 ppm at time of study	last 20 years; levels not specified	hip, wrist, and humerus	exposure: 2,563		
Chachra *et al.* (2010)	Ecological or correlational study Canada	Patients undergoing total hip arthroplasty in one hospital in each region in	Hypothesized that the direct measurement of bone tissue from	Fluoridated region (Toronto), 1 ppm; fluoridation in	Non-fluoridated region (Montreal), levels not specified	Properties of bone samples: fluoride content	Toronto: 53 (27 female); Montreal: 39 (24 female)	Toronto 66 and Montreal 70	Toronto (fluoridated) 50.9; Montreal (non-

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		Canada between September 1996 and August 2000	individuals residing in municipalities with and without fluoridated water would reveal a relationship between the fluoride content and the structural or mechanical properties of bone.	place since 1963; duration of exposure not specified		(including density) and structural or mechanical properties of bone			fluoridated) 61.5
O'Sullivan and O'Connell (2014)	Cross-sectional survey; water fluoridation status assigned on ecological/population basis Ireland	Nationally representative sample of people aged 50 years or over (and their spouses and partners of any age) resident in Ireland	Matched data from 'The Irish Longitudinal Study on Ageing (TILDA)' with Census 2006 data on the type of water supply in the local area to assess the relationships between water fluoridation and oral health and bone density in older adults.	Percentage of households with fluoridated water supply in electoral area, specific levels not specified. Each local authority was responsible for fluoridating its own water supply within the range of 0.6–0.8 ppm between 2002 and time of data collection. Prior to 2002, the	Fluoride levels in non-fluoridated areas not specified, generally not more than 0.3 ppm in Ireland. According to the 2006 Census, around 84% of households have fluoridated water supplies, which is unsurprising given that all the main urban areas receive	BMD and body mass index	4,977 people aged 50 years and over	Not reported	51

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
				target range for fluoridation was 0.8–1.0 ppm. Fluoridation status was based on 2006 Census address data and 2006 local government water source data; duration not specified; fluoridation started in 1964.	local government water supplies.				
Osteoporosis									
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Investigated the effect of drinking community fluoridated tap water by comparing the incidence of hip fractures, osteoporosis, and bone cancer prevalence in Cheongju, South Korea, where the area was naturally divided	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in 11 areas, withdrawn in 2004 in all areas	7 areas that did not receive CWF.	Osteoporosis	CWF: 4,406,021; non-CWF: 2,270,959	Not reported	CWF 50.1; Non-CWF 50.4

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			depending on the implementation status of drinking community fluoridated water.						

Table 57 Study characteristics for studies examining fractures

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Fractures									
Danielson *et al.* (1992)	Ecological or correlational study USA	Cases of hip fractures in those aged 65 years or over requiring hospitalisation 1984–1990 in rural communities in Utah, USA (records of Medicare admissions and discharges) (excluding those aged under 65 years; surgical revision of hip fracture; cases where	Examined the standardised age-specific rates of hip fracture in adults aged 65 years and over in a community with a water supply fluoridated to 1 ppm and two communities without a fluoridated water supply in Utah.	One rural community in Utah, USA, with fluoridated water (1 ppm) since 1966; duration of exposure: 24 years	Two rural communities in Utah, USA, with non-fluoridated water (<0.3 ppm)	Hip fracture	Fluoridated community: 84; non-fluoridated community: 162	Not reported	Fluoridated community 77.4; non-fluoridated community 80.3

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		fracture is possibly due to metastatic cancer; or represented a second fracture); age-specific populations for those aged 65 years or over drawn from the Utah Peer Review Organisation files of Medicare recipients, obtained from annual census counts carried out by the Social Security Administration							
Jacobsen *et al.* (1992)	Ecological or correlational study USA	Cases of hip fracture recorded by the Health Care Financing Administration and the United States Department of Veterans Affairs for white	Examined the association of water fluoridation practices in the USA and the incidence of hip fracture (rates were standardised and adjusted).	Counties that were >50% urban; natural fluoride level was <0.3 ppm; <10% of the population was served with fluoridated water prior to change, which	Counties that were >50% urban; <10% of the population was served with fluoridated water; natural fluoride levels were <0.3 ppm during 1985	Hip fracture	Not reported	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		women and men aged 65 years and over for the period 1984–1987; excluding those aged under 65 years, of non-white race, located in Puerto Rico or missing a ZIP code, cases of second fracture of hip, cases where fracture was secondary to metastatic or primary neoplastic disease, or cases where primary discharge diagnosis was for late effects of hip fracture or orthopaedic aftercare		increased to >67% of the population served with fluoridated water within a period of 3 years; duration of exposure not specified					
Suarez-Almazor *et al.* (1993)	Ecological or correlational study Canada	Cases of hip fracture (discharge diagnosis) in individuals aged 45 years or over	Compared the hip fracture hospital separation rates of two cities with and	Hip fracture admissions in Edmonton, where water has been fluoridated to 1.0 mg/L since	Hip fracture admissions in Calgary, where water is not fluoridated, and natural levels are	Hip fracture	Primary hip fractures in Edmonton: 2,479; primary hip fractures in	Not reported	Edmonton (fluoridated) 69.0; Calgary (non-fluoridated) 73.1

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		living in Edmonton or Calgary, Alberta, Canada, who were admitted to hospital in Alberta between 1981 and 1987; population estimates from the Alberta Bureau of Statistics with linear interpolation	without CWF in the Canadian province of Alberta for people aged 45 years or over.	1967; duration of exposure not specified	on average about 0.3 mg/L		Calgary: 2,392		
Jacobsen *et al.* (1993)	Ecological or correlational study USA	All incident hip fractures occurring in Rochester, Minnesota, USA from 1950 to 1969 among persons aged 50 years and over. Data taken from Mayo Clinic master index of all diagnoses and surgical procedures.	Determined the incidence of hip fracture among men and women aged 50 years and older for the 10 years prior to and the 10 years following the fluoridation of the public water supply in Rochester, Minnesota.	Rochester, Minnesota, USA, where water fluoridation was introduced in 1960, with levels maintained at 1.1 ppm; duration of exposure: >10 years	Rochester, Minnesota, USA, prior to introduction of water fluoridation in 1960	Hip fracture	651 (383 women)	Not reported	58.8

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Kröger *et al.* (1994)	Cross-sectional survey Finland	Kuopio Osteoporosis Study: all perimenopausal women aged 47–56 years residing in Kuopio Province, Finland in February 1989; random stratified sample of those willing to undergo bone densitometry	Investigated the influence of fluoridated drinking water on axial bone mineral density and the occurrence of fractures in a large perimenopausal population.	Women using fluoridated drinking water for more than 10 years (1.0–1.2 mg/L)	Women who did not have access to artificially fluoridated drinking water or who had used it for less than 10 years. The fluoride content of drinking water is low (0.0–0.3 mg/L).	Incidence of wrist fractures, ankle fractures, other fractures, and all fractures	Fluoride group: 969; non-fluoride group: 2,253	Fluoride group 53.2, non-fluoride group 53.5	100
Cauley *et al.* (1995) (Pittsburgh) ●	Cross-sectional survey USA	Women aged 65 years or over, excluding black women (due to reduced incidence of hip fractures) and women unable to walk without the assistance of another person or who had bilateral hip replacements; most recruited from voter registration lists	Evaluated whether women aged 65 years or over with greater lifetime exposure to residential fluoridated water have higher axial and appendicular bone mass and suffer fewer fractures than women compared with a lower lifetime	Years of exposure to fluoridated community water supplies recorded for each participant; exposure duration range: 1–38 years; mean fluoride concentration 1.01 ppm (±0.21 SD) for fluoridated public water	Zero years of exposure to fluoridated community water supplies recorded for each participant	Bone mineral content and density for the spine and hip and at the midpoint and ultradistal radius and calcaneus. Spinal and non-spinal fractures were also recorded.	Zero years of fluoride exposure: 1,248; 1–10 years of fluoride exposure: 438; 11–20 years of fluoride exposure: 198; and >20 years of fluoride exposure: 192	Zero years fluoride exposure 70.8, 1-10 years fluoride exposure 71.2, 11-20 years fluoride exposure 70.7, and >20 years fluoride exposure 71.6	100

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		for ZIP codes within 25-mile radius of Monessen, Pennsylvania, USA	exposure to residential fluoridated water.						
Lehmann *et al.* (1998)	Ecological or correlational study Germany	Cases: patients aged 35 years or over with hip fractures admitted to local hospitals from 1987 to 1989, excluding those admitted after trauma or pathological fractures. Population: residents of two communities in Germany – capitals of districts, industrial centres of respective regions, population estimates by 5-year age groups between 1987 and 1989 obtained from	Investigated the effect of drinking water fluoridation on hip fracture incidence by comparing two similar communities (of adults aged 35 years or over) in the former German Democratic Republic, one with artificially fluoridated water at a level of 1 mg/L and one without.	Chemnitz, Germany, where water was fluoridated to 1 mg/L over a period of 30 years; duration of exposure: >10 years	Halle, Germany, where water was not fluoridated and has a naturally occurring fluoride level of 0.08–0.36 mg/L	Hip fracture	612 in Chemnitz, Germany, and 640 in Halle, Germany	Not reported	79.9 Chemnitz (fluoridated); 81.1 Halle (non-fluoridated)

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		each city's Bureau of Statistics.							
Phipps *et al.* (2000) (Pittsburgh) ●	Cross-sectional survey USA	Prospective sample of 9,704 white women aged 65 years or over in Portland, Oregon; Minneapolis, Minnesota; Baltimore, Maryland; and the Monongahela Valley, Pennsylvania, USA, recruited from jury selection and voter registration, motor vehicle records, and membership records of health plans. Excluded white women unable to walk without assistance and women who had bilateral hip	Determined whether older women with long term exposure to fluoridated water had different bone mass and rates of fracture compared with women with no exposure.	Women exposed to fluoridated water continuously for the last 20 years; levels not specified, USA standard target fluoride level was 0.7–1.2 ppm at time of study	Women with no exposure to fluoridated water for the last 20 years; levels not specified	BMD and fractures of the vertebrae, hip, wrist, and humerus	No exposure: 3,218; continuous exposure: 2,563	No exposure 74.5, Continuous exposure 73.9	100

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		replacement. Recruitment took place from 1986 to 1988.							
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Monitored the health effects of water fluoridation in England. Compared rates of hip fracture between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	Fluoridated areas, aims to fluoridate to 1 ppm; duration of exposure not specified	Non-fluoridated areas, levels not specified	Hip fracture	Population: Fluoridated 37,971,918, non-fluoridated 274,884,530 Cases: Fluoridated 45,219, non-fluoridated 303,848	Not reported	Not reported
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Investigated the effect of drinking community fluoridated tap water by comparing the incidence of hip fractures, osteoporosis, and bone cancer prevalence in Cheongju,	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in 11 areas, withdrawn in 2004 in all areas	7 areas did not receive CWF.	Hip fracture	CWF: 4,406,021; non-CWF: 2,270,959	Not reported	CWF 50.1 and Non-CWF 50.4

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			South Korea, where the area was naturally divided depending on the implementation status of drinking community fluoridated water.						

Table 58 Study characteristics for studies examining neuropsychological outcomes

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
IQ/neuropsychological development									
Broadbent *et al.* (2015)	Prospective cohort study New Zealand	Representative sample of children born between 1 April 1972 and 31 March 1973 in Dunedin, New Zealand	Tested the hypothesis that spending childhood in an area with CWF was associated with lower IQ in childhood and adulthood.	Residence in area with CWF (0.7–1.0 ppm), use of 0.5 mg fluoride tablets (ever/never), and/or use of fluoridated toothpaste (always/sometimes/never/unknown) by the age of 5 years; duration of exposure: up to	Residence in area without CWF (0.0–0.3 ppm), use of 0.5 mg fluoride tablets (ever/never), and/or use of fluoridated toothpaste (always/sometimes/never/unknown) by the age of 5 years	IQ	992 (childhood IQ); 942 (adult IQ)	38	48

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
				5 years (preschool years only evaluated)					
Green *et al.* (2019) (MIREC) 🔴	Prospective cohort study Canada	Pregnant women from 10 cities in Canada, who could communicate in English/French, aged over 18 years, within the first 14 weeks of pregnancy; excluded if there was a known foetal abnormality, medical complications, or illicit drug use during pregnancy; subset of children recruited from 6 cities	Examined the association between fluoride exposure during pregnancy and IQ scores in a prospective birth cohort.	Fluoridated water (0.59 mg/L (±0.08 SD)); duration of exposure: 9 months (prenatally) MUF concentration in fluoridated regions averaged across all three trimesters, adjusted for specific gravity 0.69 mg/L (±0.42 SD)	Non-fluoridated water (0.13 mg/L (±0.06 SD)) MUF concentration in non-fluoridated regions averaged across all three trimesters, adjusted for specific gravity 0.40 mg/L (±0.27 SD)	IQ	512 mother–child pairs with urinary fluoride, IQ, and complete covariate data; 400 mother–child pairs with fluoride intake, IQ, and complete covariate data (non-fluoridated: 238; fluoridated: 162)	Mothers: Non-fluoridated 32.6 and fluoridated 32.5 Children: Non-fluoridated 3.4 and fluoridated 3.5	Children: Non-fluoridated 50 and fluoridated 51
Till *et al.* (2020) (MIREC) 🔴	Prospective cohort study Canada	2,001 pregnant women from 10 Canadian cities who could communicate in English or French, were aged over 17 years, and were <14 weeks' gestation;	Examined the association between fluoride exposure in infancy and intellectual ability in children who lived in fluoridated or	Fluoridated water (0.58 mg/L (±0.08 SD) (breastfed group) or 0.59 mg/L (±0.07 SD) (formula-fed group)); duration of exposure: 30–	0.13 mg/L (±0.06 SD) (breastfed group) or 0.13 mg/L (±0.05 SD) (formula-fed group)	IQ	Total: 398; breastfed: 200 (fluoridated: 83; non-fluoridated: 117); formula-fed: 198 (fluoridated:	Years of age at IQ testing: breastfed fluoridated 3.48 (±0.29 SD), non-fluoridated 3.34 (±0.31 SD); Formula fed:	Female sex: breastfed: fluoridated 51 non-fluoridated 53; formula

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		excluded if there was a known foetal abnormality or illicit drug use in pregnancy, or if they had any medical complications; 610 children recruited to participate in the developmental follow-up phase of the study from 6 of the cities in the original cohort	non-fluoridated cities in Canada	48 months postnatally			68; non-fluoridated: 130)	fluoridated 3.53 (±0.28 SD), non-fluoridated 3.37 (±0.3 SD)	fed: fluoridated 54, non-fluoridated 47.
Farmus *et al.* (2021) (MIREC) 🔴	Prospective cohort study Canada	MIREC longitudinal cohort: women in Canadian cities aged 18 years or over, at less than 14 weeks' gestation, who spoke English or French. Exclusion criteria included foetal abnormalities, medical complications, illicit drug use	Examined the impact of fluoride exposure on children's IQ scores as a function of exposure timing and sex in the same cohort	44% of pairs resident in fluoridated cities; no information on dose or duration of exposure	56% of pairs resident in non-fluoridated cities; no information on duration	IQ	596 mother–child pairs with fluoride intake, IQ, and complete demographic and covariate data (non-fluoridated: ~334; fluoridated: ~262)	Mothers: 32.4 (SD=5.1) whole sample Children: 3.4 (SD = 0.3) whole sample	Children: 51.1

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		during pregnancy, or other details previously described. Sample for this study: 601 mother–child dyads from follow-up phase (MIREC Child Development Plus); data from 5 mother–child dyads were excluded due to the mothers' declining prenatal and birth data collection (i.e. trimester fluoride exposures, demographic information, covariates, and offspring date of birth), leaving 596 mother–child dyads for full analytic sample.							
Ibarluzea et al. (2021)	Prospective cohort study Spain	Mother-child pairs recruited during first trimester of pregnancy from	To assess the association between maternal fluoride exposure during	Areas with artificially fluoridated water, fluoride concentration	Areas without artificially fluoridated water, fluoride	Neuropsychological development of children at 1 year (Bayley	393 women with complete information, 316 children	1-year follow up: Mothers 31.2 (3.4) years,	1-year follow up: 53.8%; 4-year

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		one referral hospital in Spain; women aged 16 years or over, with singleton pregnancy achieved without assisted reproduction techniques, planned birth in the referral hospital, no communication problems in Spanish or Basque	pregnancy and neurodevelopmental outcomes in 1- and 4-year-old children"	CWF areas 0.81 (±0.15 SD) mg/L	concentration <0.1 mg/L	Scales of Infant Development) and 4 years (McCarthy Scales of Children's Abilities)	included at age 1, 248 children included at age 4	children 14.6(0.8) months; 4-year follow up: Mothers 31.5(3.4) years, children 4.4(0.1) years	follow up: 49.6%
Neurodevelopmental disorders									
Barberio *et al.* (2017a) (CHMS) ●	Cross-sectional survey Canada	Population-based sample of Canadian children aged 3–12 years living in private households in the 10 provinces (subset for whom information on sources of fluoride exposure was available)	Examined the association between fluoride exposure and reported diagnosis of a learning disability among a population-based sample of Canadian children aged 3–12 years.	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3 of surveys. Spot urine samples were available for a subsample of the	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3 of surveys. Spot urine samples were available for a subsample of the	Learning disability diagnosis (attention deficit disorder, no hyperactivity/ADHD/dyslexia/other)	Fluoride subsample (Cycle 2: 1,120; Cycle 3: 1,101)	Cycle 2: 6.8 and cycle 3: 7.03 years	Cycle 2: 50.6 and cycle 3: 49.2

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
				respondents for Cycles 2 and 3,as specific gravity-adjusted urinary fluoride (micromoles per litre (µmol/L)) and creatinine-adjusted urinary fluoride (micromoles per millimole (µmol/mmol)); duration of exposure not specified	respondents for Cycles 2 and 3, as specific gravity-adjusted urinary fluoride (µmol/L) and creatinine-adjusted urinary fluoride (µmol/mmol)				
Riddell *et al.* (2019) (CHMS)	Cross-sectional survey Canada	National sample of Canadian youth aged 6–17 years from a population-based sample of Canadian residents aged 3–79 years living in private households (subset for whom information on sources of fluoride exposure was available (i.e. urine samples and	Examined the relationship between urinary and tap water fluoride concentrations and attention-related outcomes in a national sample of Canadian youth aged 6 to 17 years.	Fluoride in urine and tap water as a continuous variable. CWF as a dichotomous variable with mixed areas excluded. Duration of exposure not specified.	Fluoride in urine and tap water as a continuous variable. CWF as a dichotomous variable with mixed areas excluded.	ADHD diagnosis and scores	Fluoride in urine (1,877) and tap water (710) as continuous variables. CWF is a dichotomous variable (total: 1,722 (932 fluoridated; 790 non-fluoridated)) with mixed areas excluded.	Fluoride in urine (11.3 years) and tap water (11.2 years) as continuous variables. CWF is a dichotomous variable (11.3 years) with mixed areas excluded.	Fluoride in urine (48.8% female) and tap water (47.3% female) as continuous variables. CWF is a dichotomous variable (49.2%

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		tap water samples))							female) with mixed areas excluded .

Table 59 Study characteristics for studies examining bone cancers

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Osteosarcoma									
Hrudey *et al.* (1990)	Ecological or correlational study Canada	Cases of osteosarcoma recorded by the Alberta Cancer Board registry, Canada	Compared the incidence of osteosarcoma in two cities, with and without CWF, in the Canadian province of Alberta.	Incidence of osteosarcoma in Edmonton, Alberta, Canada, where water has been fluoridated to 1.0 mg/L since 1967, recorded by the Alberta Cancer Board	Incidence of osteosarcoma recorded by the Alberta Cancer Board in Calgary, where water is not fluoridated (natural levels average 0.3 mg/L)	Osteosarcoma	Edmonton (fluoridated) 26 cases, Calgary (non-fluoridated) 29 cases	Not reported	Edmonton (fluoridated) 57.7, Calgary (non-fluoridated) 44.8
Mahoney *et al.* (1991)	Ecological or correlational study USA	Bone cancers recorded by the New York State Cancer Registry, USA, primary malignancies of the bone	Compared average annual bone cancer and osteosarcoma incidence rates by grouped years 1975 to 1987 in fluoridated and	Primary analysis: fluoridated areas in New York State, exclusive of New York City (due to lack of outcome data before 1973). Additional	Primary analysis: non-fluoridated areas in New York State, including some metropolitan areas that have maintained	Bone cancer, including osteosarcoma	Bone cancer cases (n=228) and osteosarcoma cases (n= 108)	Not reported	Bone cancer 40.9 and osteosarcoma 46.3

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			non-fluoridated areas.	analysis: fluoridated counties located within standard metropolitan statistical areas (i.e. urbanised areas) and fluoridated counties not located within standard metropolitan statistical areas. Level of fluoridation not specified, and duration of exposure not specified.	non-fluoridated water supplies. Additional analysis: three urbanised metropolitan areas that have maintained non-fluoridated water supplies.				
McGuire *et al.* (1991)	Matched case-control study USA	Cases: patients diagnosed with osteosarcoma between 1980 and 1990, aged under 40 years at diagnosis, identified from the University of Iowa Cancer Registry and the medical records of the Division of Orthopaedics, St 'Joseph's	Tested the hypothesis that fluoride is a risk factor for osteosarcoma, this case-control study compared the complete residential fluoride histories of osteosarcoma patients with matched hospital-based controls of	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	22 cases	Not reported	40.9

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		Hospital in Omaha, Nebraska, USA, excluding patients with prediagnosis history of radiation therapy or kidney disease. Hospital-based controls from orthopaedics department matched by age, gender, and county of residence at time of diagnosis.	similar age, gender, and county of residence.						
Cohn (1992)	Ecological or correlational study USA	Cases of osteosarcoma in those aged under 20 years in seven central New Jersey, USA counties, compiled from the New Jersey Cancer Registry between 1979 and 1987; population data gathered from 1980 United	Compared osteosarcoma incidence between 1979 and 1987 in young people under 20 years of age living in New Jersey in artificially fluoridated areas with non-fluoridated areas.	Fluoridated municipalities where >85% of the population was supplied with fluoridated water from at least the early 1970s to at least 1987; USA standard target fluoride level was 0.7–1.2 ppm at time of study	Non-fluoridated municipalities where <10% of the population was supplied with fluoridated water; levels not specified	Osteosarcoma	Under 20 years of age: Population: 721,347 Cases: 30	Not reported	45.5

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		States Census on a municipality level							
Bassin *et al.* (2006) (Harvard) 🔵	Case-control study USA	Cases of osteosarcoma diagnosed before the age of 20 years in 11 hospitals across the USA (excluding those aged over 40 years, or with a history of radiation therapy or renal dialysis); controls were patients of the same hospitals' orthopaedics departments seen within ±6 months of cases' diagnosis and matched for age (±5 years), gender, distance from hospital, same exclusion criteria	Explored age-specific and sex-specific effects and evaluated exposure to fluoride in drinking water from birth through early adolescence, limited to males and females under 20 years.	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	Cases 103 and Controls 215	Cases 13.7 and Controls 14.5	Cases 42 and Controls 43
Comber *et al.* (2011)	Ecological or correlational study	Cases of osteosarcoma recorded in the Northern Ireland	Compared the incidence of osteosarcoma in Northern Ireland	Fluoridated regions (non-rural Republic of Ireland electoral	Non-fluoridated regions (rural Republic of	Osteosarcoma	183 cases	Not reported	37.7

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
	Ireland	Cancer Registry and the National Cancer Registry of Ireland between 1994 and 2006. Population denominator assigned by electoral district – mid-year population estimates for Northern Ireland, provided annually by the Northern Ireland Statistics and Research Agency; Republic of Ireland census data for 1996, 2002, and 2006 interpolated for intervening years.	with that in the Republic of Ireland to establish if differences in incidence between the two regions could be related to their different drinking water fluoridation policies.	divisions), levels not specified, targets 0.6 – 1.0 ppm during this time period; duration not specified; water fluoridation signed into law in 1960	Ireland electoral divisions and Northern Ireland); levels not specified, generally not more than 0.3 ppm in Ireland				
National Fluoridation Information Service (2013)	Ecological or correlational study New Zealand	Cases of osteosarcoma recorded in the New Zealand Cancer Registry between 2000 and 2008; rate	Examined risk of the bone cancer osteosarcoma by CWF status.	Census area units served by CWF; levels not specified but generally fluoridated to 0.7–1.0 ppm in	Census area units not served by CWF; levels not specified, generally not more than 0.3	Osteosarcoma	127 cases (fluoridated 58 and non-fluoridated 69)	Not reported	38.6

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		per 1,000,000 population calculated but no information given on source of denominator data		New Zealand; duration of exposure not specified	ppm in New Zealand				
Blakey *et al.* (2014)	Ecological or correlational study United Kingdom	Cases of osteosarcoma or Ewing sarcoma diagnosed in Great Britain from 1980 to 2005 drawn from population-based cancer registries. Denominator data derived from national decennial census data, population counts from previous censuses adjusted to be compatible with 2001 Census boundaries, census wards in England and Wales, and postcode	Examined whether increased risk of primary bone cancer was associated with living in areas with higher concentrations of fluoride in drinking water.	Level of fluoride in water for census small area units in 2001; duration of exposure not specified	Level of fluoride in water for census small area units in 2001 (introduction data: optimal fluoride: 0.7–1.2 ppm; non-fluoridated: 0.3 ppm (confers no dental benefit))	Osteosarcoma and Ewing sarcoma	Osteosarcoma cases 2,566 and Ewing sarcoma cases 1,650	Not reported	Osteosarcoma 41.8 and Ewing sarcoma 40.1

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		sectors in Scotland.							
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Monitored the health effects of water fluoridation arrangements in England. Compared rates of osteosarcoma between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Osteosarcoma	Population under 25 years: fluoridated areas 31,313,151, non-fluoridated areas 216,921,400 Cases under 25 years: fluoridated areas 148, non-fluoridated areas 949 Population 50 years and over: fluoridated areas 33,080,465, non-fluoridated areas 232,282,090 Cases 50 years and over: fluoridated areas 73, non-fluoridated areas 587.	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Kim *et al.* (2020) (Harvard) 🔵	Case-control study USA	Cases: Phase 1: all patients diagnosed with osteosarcoma by participating departments in the USA from 1989 to 1993, histologically confirmed; Phase 2: cases of osteosarcoma identified by physicians in participating departments, diagnosed and treated with primary osteosarcoma confirmed by surgical pathology reports from 1994 to 2000. Controls: Phase 1: patients of record from 1989 to 1993 with other bone tumours or nonneoplastic conditions identified from same	Assessed whether living in a fluoridated community in the USA is a risk factor for osteosarcoma; This includes all participants of the Bassin *et al.* 2006.	Estimated level of fluoride in drinking water at each address of residence; lifetime exposure	Estimated level of fluoride in drinking water at each address of residence	Osteosarcoma	All 645; Cases 236; Controls 409;	Not reported	All 39.5; Cases 39.8; Controls 39.4

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		orthopaedic surgery departments as cases; Phase 2: patients from same departments as cases from 1994 to 2000 with newly diagnosed malignant bone tumours other than osteosarcoma (tumour controls) and benign tumours and nonneoplastic conditions (orthopaedic controls), including ambulatory orthopaedic patients.							
Ewing sarcoma									
Blakey *et al.* (2014)	Ecological or correlational study United Kingdom	Cases of osteosarcoma or Ewing sarcoma diagnosed in Great Britain from 1980 to 2005 drawn	Examined whether increased risk of primary bone cancer was associated with living in areas	Level of fluoride in water for census small area units in 2001; duration of exposure not specified	Level of fluoride in water for census small area units in 2001 (introduction	Osteosarcoma and Ewing sarcoma	Osteosarcoma cases 2,566 and Ewing sarcoma cases 1,650	Not reported	Osteosarcoma 41.8 and Ewing sarcoma 40.1

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		from population-based cancer registries. Denominator data derived from national decennial census data, population counts from previous censuses adjusted to be compatible with 2001 Census boundaries, census wards in England and Wales, and postcode sectors in Scotland.	with higher concentrations of fluoride in drinking water.		data: optimal fluoride: 0.7–1.2 ppm; non-fluoridated: 0.3 ppm (confers no dental benefit))				
Bone cancer									
Mahoney *et al.* (1991)	Ecological or correlational study USA	Bone cancers recorded by the New York State Cancer Registry, USA, primary malignancies of the bone	Compared average annual bone cancer and osteosarcoma incidence rates by grouped years 1975 to 1987 in fluoridated and	Primary analysis: fluoridated areas in New York State, excluding New York City (due to lack of outcome data before 1973). Additional	Primary analysis: non-fluoridated areas in New York State, including some metropolitan areas that have maintained	Bone cancer, including osteosarcoma	Bone cancer cases (n=228) and osteosarcoma cases (n= 108)	Not reported	Bone cancer 40.9 and osteosarcoma 46.3

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			non-fluoridated areas.	analysis: fluoridated counties located within standard metropolitan statistical areas (i.e. urbanised areas) and fluoridated counties not located within standard metropolitan statistical areas. Level of fluoridation not specified; duration of exposure not specified.	non-fluoridated water supplies. Additional analysis: three urbanised metropolitan areas that have maintained non-fluoridated water supplies.				
Lee *et al.* (2020)	Ecological or correlational study South Korea	Population: residents of Cheongju region, South Korea. Cases: cases of hip fracture, osteoporosis, and bone cancer identified from National Health Insurance Service data.	Investigated the effect of drinking community fluoridated tap water by comparing the incidence of hip fractures, osteoporosis, and bone cancer prevalence in Cheongju, South Korea, where the area was	Fluoridated areas: dose not specified; duration of exposure not specified; CWF introduced in 1982 in 10 areas and in 1997 in 11 areas, withdrawn in 2004 in all areas	7 areas did not receive CWF.	Bone cancer	Populations: CWF 4,406,021 and Non-CWF 2,270,959	Not reported	CWF 50.1 and Non-CWF 50.4

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			naturally divided depending on the implementation status of drinking community fluoridated water.						

Table 60 Study characteristics for studies examining other cancers

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Secondary bone cancer									
Crnosija *et al.* (2019)	Ecological or correlational study USA	Population: inpatient cancer patients drawn from the Statewide Planning and Research Cooperative System database run by the New York State Department of Health, USA; excluded those without a complete ZIP code, a patient ID code, or a New York State residency status, or who were aged under 18 years. Cases: presence or absence of the diagnosis code for secondary bone cancer (metastasis).	Ascertained whether county-level rates of bone metastasis vary by percentage of population drinking fluoridated water in that county.	Limited variation in concentration of fluoride in fluoridated water supplies (45 counties received 0.7 mg/L; 2 counties 0.8 mg/L; 1 county 0.5 mg/L; and 1 county 0.4 mg/L); fluoridation was therefore evaluated by percentage of the population in each county that received public fluoridated water, and divided into three categories: <25%, 25–75%, and >75%; duration not specified.	Limited variation in concentration of fluoride in fluoridated water supplies (45 counties received 0.7 mg/L; 2 counties 0.8 mg/L; 1 county 0.5 mg/L; and 1 county 0.4 mg/L); fluoridation was therefore evaluated by percentage of the population in each county that received public fluoridated water, and divided into three categories: <25%, 25–75%, and >75%; duration not specified.	Secondary bone cancer	190,636	Not reported	Not reported
Bladder cancer									
Young *et al.* (2015)	Ecological or correlational study	Adults and children in England	Monitored the health effects of water	Fluoridated areas, target level 1.0 ppm; duration of	Non-fluoridated areas	Bladder cancer	Population: fluoridated areas 67,978,298,	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
	England		fluoridation arrangements in England. Compared rates of bladder cancer between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	exposure not specified			non-fluoridated areas 487,149,150 Cases: fluoridated areas: 11,327; non-fluoridated areas: 84,780		
All cancers									
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Monitored the health effects of water fluoridation arrangements in England. Compared rates of all cancers between areas	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	All cancers	Population: fluoridated areas 25,314,612, non-fluoridated areas 183,256,350 Cases: fluoridated areas: 131,288	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
			according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).				non-fluoridated areas: 921,583		

Table 61 Study characteristics for studies examining endocrine conditions

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Thyroid functioning									
Peckham *et al.* (2015)	Ecological or correlational study England	Cases of hypothyroidism recorded by general practices in England in 2012 and 2013 using the Quality and Outcomes Framework	Examined differences in prevalence of hypothyroidism between fluoridated and non-fluoridated areas	Levels of fluoride concentration assigned to general practices based on postcodes; target level 1 ppm (2012 data); maximum level >0.3 mg/L for all	Levels of fluoride concentration assigned to general practices using postcodes; target level not specified (2012 data); maximum level ≤0.3 mg/L for all general	Hypothyroidism	946 general practices recruited; number included in CWF analysis not clear	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
				practices in the West Midlands; duration: <1 year	practices in Greater Manchester				
Barberio *et al.* (2017b) (CHMS) ●	Cross-sectional survey Canada	Population-based sample of Canadian residents aged 3–79 years living in the 10 provinces; excluded those in the 3 territories (Cycle 3 only), those who live on reserves and other Aboriginal settlements in the provinces, institutionalised residents, full-time members of the Canadian Armed Forces, and residents of certain remote regions (subset for whom	Examined the association between fluoride exposure and (1) diagnosis of a thyroid condition and (2) indicators of thyroid functioning among a national population-based sample of Canadians ((TSH aged 3–79 years, Thyroid diagnosis aged 12–79 years)	Estimates of the fluoride concentration of tap water samples (mg/L) collected at respondents' homes were available for Cycle 3; spot urine samples were available for a subsample of the respondents for Cycles 2 and 3; duration of exposure not specified	Individual-level fluoride from urine and tap water samples	Thyroid functioning and/or diagnosis	Cycle 2 – fluoride urine subsample: 2,530; Cycle 3 – fluoride urine subsample: 2,671	Cycle 2 – fluoride urine subsample: 32.55 years; Cycle 3 – fluoride urine subsample: 35.69 years	Cycle 2 – fluoride urine subsample: 44.9; Cycle 3 – fluoride urine subsample: 48.7

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		information on sources of fluoride exposure was available)							
Malin *et al.* (2018) (CHMS)	Cross-sectional survey Canada	Participants in the CHMS: randomly selected individuals aged 3–79 years living in 16 sites across the 10 Canadian provinces; excluded residents of the 3 territories, reserves, and Aboriginal settlements; full-time members of the Canadian Armed Forces; institutionalised individuals; and those living in remote areas. Sample restricted to participants	Examined whether the relationship between fluoride exposure and thyroid function is modified by iodine status among adults participating in a Canadian population-based survey	Urinary fluoride concentrations measured in spot samples using an ion-selective electrode and adjusted for specific gravity; duration of exposure not specified	Urinary fluoride concentrations were measured in spot samples using an ion-selective electrode and adjusted for specific gravity	TSH levels	Approximately 1,000, representing 6,914,124 adults in Canada aged 18–79 years	46.5	48.5

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		aged over 18 years and those who provided urine/tap water samples, excluding pregnant women and those with iodine levels above 2.37 µmol/L.							
Sleep disturbance (pineal gland functioning)									
Cunningham *et al.* (2021) (CHMS) 🟢	Cross-sectional survey Canada	Participants in the CHMS: randomly selected individuals aged 3–79 years living in Canada; excluded residents of reserves and Aboriginal settlements; full-time members of the Canadian Armed Forces; institutionalised individuals; and	Examined the association between fluoride exposure and sleep outcomes in a large Canadian sample (aged 16 –79 years) using cross-sectional data from Cycle 3 (2012–2013) of the Canadian Health Measures Survey (CHMS).	Fluoridated (missing or mixed fluoridation data analysed separately); duration of exposure not specified	Not fluoridated	Sleep disturbance due to reduced functioning of pineal gland	Urinary fluoride sample: 1,303; water fluoride sample: 1,016	Urinary fluoride sample, mean age: 42.4; Water fluoride sample, mean age: 43.3.	Urinary fluoride sample, female 51.3; Water fluoride sample, female 50.7.

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
		those living in remote areas or northern territories. Sample restricted to participants aged over 15 years and those who provided urine/tap water samples.							

Table 62 Study characteristics for studies examining renal conditions

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Renal calculi (kidney stones)									
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Monitored the health effects of water fluoridation arrangements in England. Compared rates of selected kidney stones between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Kidney stones	Population: fluoridated areas 37,971,918, non-fluoridated areas 274,884,530 Cases: fluoridated areas: 18,579; non-	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
							fluoridated areas: 141,963		

Table 63 Study characteristics for studies examining birth or birthing abnormalities

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Lowry *et al.* (2003)	Ecological or correlational study England	Mothers and their newborn children in northern England	Investigated if water fluoridation had any influence on the rate of congenital abnormalities or stillbirths in the north-east of England	Artificially fluoridated areas in northern region of England (>0.9 ppm); duration of exposure not specified	Non-fluoridated areas in northern region of England (<0.3 ppm)	Stillbirths and congenital abnormalities (all trisomies, Down syndrome, neural tube defects, and clefts)	Not reported	Not reported	Not reported
Young *et al.* (2015)	Ecological or correlational study England	Mothers and their newborn children in England	Monitored the health effects of water fluoridation arrangements in England. Compared incidence of Down syndrome between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	Fluoridated areas, aims to fluoridate to 1 ppm; duration of exposure not specified	Non-fluoridated areas	Down syndrome	Population (live births): fluoridated areas 303,818, non-fluoridated areas 2,423,482 Cases: fluoridated areas 658; non-fluoridated areas 5,961	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
Zhang *et al.* (2019)	Ecological or correlational study USA	Mothers and their newborn children in Massachusetts, USA	Examined the association of maternal dental cleaning during pregnancy and exposure to CWF on preterm birth using a population-based survey of women who recently gave birth. This study used the Massachusetts Pregnancy Risk Assessment Monitoring System data from 2009 to 2016.	Exposure to CWF (with or without dental cleaning); duration of exposure not specified	Not exposed to CWF (with or without dental cleaning)	Preterm birth or birth prior to 37 weeks' completed gestation	9,234	Not reported	100

Table 64 Study characteristics for studies examining infant abnormalities

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
SIDS									
Dick *et al.* (1999)	Case-control study with ecological assignment of CWF status	New Zealand Cot Death study covered 78% of births in New Zealand between November	Determined whether exposure to fluoridated water supplies prenatally or postnatally at the time of death or censoring was	Fluoridated areas (designated >80% (average pre- and postnatal exposure) of residences fluoridated); target level 1.0 ppm;	Non-fluoridated areas (designated <20% of residences fluoridated	Cases of SIDS	Total: 2,285 (485 SIDS cases; 1,800 controls)	Not reported	Not reported

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
	New Zealand	1987 and October 1990	associated with a higher risk of SIDS.	duration of exposure up to 1 year; partially fluoridated areas (20–80% of residences fluoridated) excluded from analysis	(average pre- and postnatal exposure))				

Table 65 Study characteristics for studies examining all-cause mortality

Paper	Study design Country	Population	Research question	Exposure	Comparator	Outcome	Sample size for analysis	Mean age	Gender (% female)
All-cause mortality									
Young *et al.* (2015)	Ecological or correlational study England	Adults and children in England	Monitored the health effects of water fluoridation arrangements in England. Compared counts of death (all-cause mortality) between areas according to whether the level of fluoride in drinking water is adjusted (fluoridated) or not (non-fluoridated).	Fluoridated areas, target level 1.0 ppm; duration of exposure not specified	Non-fluoridated areas	Count of deaths	Population: fluoridated areas 25,314,612, non-fluoridated areas 183,256,350 Cases: fluoridated areas: 233,922; non-fluoridated areas: 1,602,206	Not reported	Not reported

Appendix M Quality assessment scores

Table 66 Quality assessment ratings for studies examining bone characteristics

Item	Kröger *et al.* (1994)	Cauley *et al.* (1995) (Pittsburgh) ●	Arnold *et al.* (1997)	Lehmann *et al.* (1998) (BMD)	Phipps *et al.* (2000) (Pittsburgh) ●	Chachra *et al.* (2010)	O'Sullivan and O'Connell (2014)	Lee *et al.* (2020)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	No	No	Not reported	Yes	Yes	Not reported	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	No	No	No	Partly	No	No	Yes
5B. Was a description of variance provided?	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No	Yes	No	No	Yes	No	No	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Item	Kröger *et al.* (1994)	Cauley *et al.* (1995) (Pittsburgh) ●	Arnold *et al.* (1997)	Lehmann *et al.* (1998) (BMD)	Phipps *et al.* (2000) (Pittsburgh) ●	Chachra *et al.* (2010)	O'Sullivan and O'Connell (2014)	Lee *et al.* (2020)
10. Was the exposure(s) assessed more than once over time?	No	No	No	No	No	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Almost all	Almost all	Almost all	Almost all	Almost all	None	Partial	Limited
Quality rating	3.0 (moderate)	2.0 (low)	2.0 (low)	3.0 (moderate)	3.5 (high)	1.0 (low)	2.5 (low)	3.0 (moderate)

Table 67 Quality assessment ratings for studies examining fractures

Item	Kröger *et al.* (1994)	Danielson *et al.* (1992)	Jacobsen *et al.* (1992)	Suarez-Almazor *et al.* (1993)	Jacobsen *et al.* (1993)	Cauley *et al.* (1995) (Pittsburgh) ●	Lehmann *et al.* (1998) (Fractures)	Phipps *et al.* (2000) (Pittsburgh) ●	Young *et al.* (2015)	Lee *et al.* (2020)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Item	Kröger *et al.* (1994)	Danielson *et al.* (1992)	Jacobsen *et al.* (1992)	Suarez-Almazor *et al.* (1993)	Jacobsen *et al.* (1993)	Cauley *et al.* (1995) (Pittsburgh) ●	Lehmann *et al.* (1998) (Fractures)	Phipps *et al.* (2000) (Pittsburgh) ●	Young *et al.* (2015)	Lee *et al.* (2020)
populations (including the same time period)?										
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes	Yes	Yes	No	Yes	Partly	Yes	Yes
5B. Was a description of variance provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	No
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No	No	No	No	No	Yes	No	Yes	No	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. Was the exposure(s) assessed more than once over time?	No	No	No	No	No	No	No	No	No	No

Item	Kröger *et al.* (1994)	Danielson *et al.* (1992)	Jacobsen *et al.* (1992)	Suarez-Almazor *et al.* (1993)	Jacobsen *et al.* (1993)	Cauley *et al.* (1995) (Pittsburgh) ●	Lehmann *et al.* (1998) (Fractures)	Phipps *et al.* (2000) (Pittsburgh) ●	Young *et al.* (2015)	Lee *et al.* (2020)
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Almost all	Partial	Partial	Partial	Partial	Almost all	Partial	Almost all	Partial	Partial
Quality rating	3.5 (high)	3.5 (high)	3.5 (high)	3.5 (high)	3.0 (moderate)	2.0 (low)	3.5 (high)	3.5 (high)	3.5 (high)	3.5 (high)

Table 68 Quality assessment ratings for studies examining neuropsychological outcomes

Item	Broadbent *et al.* (2015)	Green *et al.* (2019) (MIREC) ●	Till *et al.* (2020) (MIREC) ●	Farmus *et al.* (2021) (MIREC) ●	Ibarluzea *et al.* (2021)	Barberio *et al.* (2017a) (CHMS) ●	Riddell *et al.* (2019) (CHMS) ●
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes	No	No	No	Not reported	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes	Yes	Yes	Yes	Not reported

Item	Broadbent *et al.* (2015)	Green *et al.* (2019) (MIREC) 🔴	Till *et al.* (2020) (MIREC) 🔴	Farmus *et al.* (2021) (MIREC) 🔴	Ibarluzea *et al.* (2021)	Barberio *et al.* (2017a) (CHMS) 🟢	Riddell *et al.* (2019) (CHMS) 🟢
5A. Was a sample size justification, power description, or variance and effect estimates provided?	No	No	No	No	No	No	No
5B. Was a description of variance provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	Yes	Yes	Yes	Yes	No	No
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
10. Was the exposure(s) assessed more than once over time?	No	No	No	No	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Yes	Not reported	Not reported	Yes	Yes	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Yes	No	No	No	No	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Almost all	Partial	Almost all	Almost all	None	Limited
Quality rating	3.0 (moderate)	2.0 (low)	1.5 (low)	2.0 (low)	2.0 (low)	2.0 (low)	2.0 (low)

Table 69 Quality assessment ratings for ecological studies examining bone cancers

Item	Hrudey *et al.* (1990)	Mahoney *et al.* (1991)	Cohn (1992)	Comber *et al.* (2011)	National Fluoridation Information Service (2013)	Blakey *et al.* (2014)	Young *et al.* (2015)	Lee *et al.* (2020)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Not reported	Yes	Not reported	Not reported	Not reported	Not reported	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
5B. Was a description of variance provided?	No	Yes	Yes	Yes	No	Yes	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Not clear	No	Yes	Yes	Yes	Yes	Yes	No
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No	No	No	No	No	Yes	No	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Item	Hrudey *et al.* (1990)	Mahoney *et al.* (1991)	Cohn (1992)	Comber *et al.* (2011)	National Fluoridation Information Service (2013)	Blakey *et al.* (2014)	Young *et al.* (2015)	Lee *et al.* (2020)
10. Was the exposure(s) assessed more than once over time?	No	No	No	No	No	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Limited	Limited	Limited	Limited	Limited	Limited
Quality rating	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)

Table 70 Quality assessment ratings for case-control studies examining osteosarcoma

Item	McGuire *et al.* (1991)	Bassin *et al.* (2006) (Harvard)	Kim *et al.* (2020) (Harvard)
1. Was the research question or objective in this paper clearly stated and appropriate?	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes
3. Was an appropriate target population clearly defined per the research question? Did the cases adequately represent the cases that arose in the target population?	Partly	Partly	Partly
4. Did the authors include a sample size justification?	No	No	No
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes	Yes	Partly; timeframe not clear
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
7. Were the cases clearly defined and differentiated from controls?	Yes	Yes	Yes
8. If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?	Not applicable	Not applicable	Not applicable
9. Was there use of concurrent controls?	Yes	Yes	No
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	Yes	Partly	Partly
11. Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (including the same time period) across all study participants?	Yes	Yes	Yes
12. Were the assessors of exposure/risk blinded to the case or control status of participants?	Not reported	Not reported	Not reported
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Partial	Partial	Partial
Quality rating	3.5 (high)	2.5 (low)	2.5 (low)

Table 71 Quality assessment ratings for ecological studies examining other cancers

Item	SECONDARY BONE CANCER Crnosija *et al.* (2019)	BLADDER CANCER Young *et al.* (2015)	ALL CANCERS Young *et al.* (2015)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	Yes
5B. Was a description of variance provided?	No	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	No	Yes	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	No	Yes	Yes
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	Yes	No	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
10. Was the exposure(s) assessed more than once over time?	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Unable to determine	Limited	Unable to determine
Quality rating	3.0 (moderate)	3.0 (moderate)	3.0 (moderate)

Table 72 Quality assessment ratings for studies examining endocrine conditions

Item	Peckham *et al.* (2015)	Barberio *et al.* (2017b) (CHMS)	Malin *et al.* (2018) (CHMS)	Cunningham *et al.* (2021) (CHMS)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	No	No	No
5B. Was a description of variance provided?	Yes	Yes	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	No	No	No
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes	No	No	No
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	Yes	Yes	Yes	Yes
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes
10. Was the exposure(s) assessed more than once over time?	No	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Partial	Unable to determine
Quality rating	3.0 (moderate)	2.0 (low)	2.5 (low)	2.0 (low)

Table 73 Quality assessment ratings for studies examining renal conditions

Item	Young *et al.* (2015)
1. Was the research question or objective in this paper clearly stated?	Yes
2. Was the study population clearly specified and defined?	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes
5B. Was a description of variance provided?	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes
10. Was the exposure(s) assessed more than once over time?	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited
Quality rating	3.0 (moderate)

Table 74 Quality assessment ratings for studies examining birth and birthing abnormalities

Item	Lowry *et al.* (2003)	Young *et al.* (2015)	Zhang *et al.* (2019)
1. Was the research question or objective in this paper clearly stated?	Yes	Yes	Yes
2. Was the study population clearly specified and defined?	Yes	Yes	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes	Yes	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes	Yes	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes	Yes	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes	Yes	No
5B. Was a description of variance provided?	Yes	Yes	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes	Yes	No
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	No	Yes	Yes
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No	No	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
10. Was the exposure(s) assessed more than once over time?	No	No	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes	Yes	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported	Not reported	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable	Not applicable	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Limited	Limited	Partial
Quality rating	3.0 (moderate)	3.0 (moderate)	2.5 (low)

Table 75 Quality assessment ratings for studies examining infant abnormalities

Item	Dick *et al.* (1999)
1. Was the research question or objective in this paper clearly stated and appropriate?	Yes
2. Was the study population clearly specified and defined?	Yes
3. Was an appropriate target population clearly defined per the research question? Did the cases adequately represent the cases that arose in the target population?	Yes
4. Did the authors include a sample size justification?	No
5. Were controls selected or recruited from the same or similar population that gave rise to the cases (including the same timeframe)?	Yes
6. Were the definitions, inclusion and exclusion criteria, algorithms or processes used to identify or select cases and controls valid, reliable, and implemented consistently across all study participants?	Yes
7. Were the cases clearly defined and differentiated from controls?	Yes
8. If less than 100 percent of eligible cases and/or controls were selected for the study, were the cases and/or controls randomly selected from those eligible?	Not reported
9. Was there use of concurrent controls?	Unable to determine
10. Were the investigators able to confirm that the exposure/risk occurred prior to the development of the condition or event that defined a participant as a case?	Partly
11. Were the measures of exposure/risk clearly defined, valid, reliable, and implemented consistently (including the same time period) across all study participants?	Partly
12. Were the assessors of exposure/risk blinded to the case or control status of participants?	Not reported
13. Were key potential confounding variables measured and adjusted statistically in the analyses? If matching was used, did the investigators account for matching during study analysis?	Limited
Quality rating	2.5 (low)

Table 76 Quality assessment ratings for studies examining all-cause mortality

Item	Young *et al.* (2015)
1. Was the research question or objective in this paper clearly stated?	Yes
2. Was the study population clearly specified and defined?	Yes
3. Was the participation rate of eligible persons at least 50%?	Yes
4A. Were all the subjects selected or recruited from the same or similar populations (including the same time period)?	Yes
4B. Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?	Yes
5A. Was a sample size justification, power description, or variance and effect estimates provided?	Yes
5B. Was a description of variance provided?	Yes
6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?	Yes
7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?	Yes
8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?	No
9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes
10. Was the exposure(s) assessed more than once over time?	No
11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?	Yes
12. Were the outcome assessors blinded to the exposure status of participants?	Not reported
13. Was loss to follow-up after baseline 20% or less?	Not applicable
14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?	Unable to determine
Quality rating	3.0 (moderate)

Appendix N Study findings

Table 77 Findings for studies examining bone characteristics

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Bone mineral density: Total body									
Arnold *et al.* (1997)		g/cm^^{2}, dual-energy X-ray absorptiometry (DXA) in array mode		ANCOVA	Mean BMD		-		
	Regina (Non-fluoridated)		24		1.044	Body weight	-	p>0.05	No difference in total body BMD between women raised in fluoridated and non-fluoridated areas
	Saskatoon (Fluoridated)		33		1.073	Body weight	-		
Bone mineral density: Lumbar spine									
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^^{3}, DXA		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.849	age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin	0.843–0.856	p<0.001	Women with continuous exposure had significantly higher BMD of the lumbar spine

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status			
	Continuous exposure for 20 years				0.871		0.865–0.877		
Bone mineral density: Lumbar spine L1–L4									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure (reference)		1243		0.842	Age		p=0.99	No association between fluoride exposure and density of the lumbar spine L1–L4
	1-10 years fluoride exposure		437		0.847	Age			
	11-20 years fluoride exposure		198		0.844	Age			
	>20 years fluoride exposure		192		0.849	Age			
Bone mineral density: Lumbar spine L2–L4									
Kröger *et al.* (1994)		g/cm^3, DXA		ANCOVA with adjustment of confounding factors	Mean BMD		-		
	All women - fluoride		969		1.151	Age, weight, menopausal status,	-	p=0.001	Perimenopausal women using

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						calcium intake, physical activity class, number of deliveries, alcohol consumption, and oestrogen use			fluoridated drinking water for more than 10 years had significantly higher BMD in the spine than those not using fluoridated water; significant differences for premenopausal women, postmenopausal women, and whole sample.
	All women - non-fluoride		2253		1.121		-		
	Premenopausal women - fluoride		281†		1.195		-	p=0.002	
	Premenopausal women - non-fluoride		701†		1.165		-		
	Postmenopausal women - fluoride		688†		1.131		-	p=0.005	
	Postmenopausal women - non-fluoride		1552†		1.101		-		
Lehmann *et al.* (1998)		g/cm^3, DXA		MANOVA/ MANCOVA	Mean BMD		Standard deviation		No significant differences in adjusted BMD between regions for

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									either men or women
	Men fluoridated		41		1.045	Age and BMI	0.171	p=0.08	
	Men non-fluoridated		98		0.997	Age and BMI	0.129		
	Women fluoridated		201		1.046	Age and weight	0.117	p=0.47	
	Women non-fluoridated		215		1.055	Age and weight	0.112		
Bone mineral density: Estimated volumetric lumbar 3									
Arnold *et al.* (1997)		g/cm^3; Volumetric estimate of BMD (bone mineral content divided by estimated volume)		ANCOVA	Mean BMD		-		
	Regina (Non-fluoridated)		24		0.216	Body weight	-	p<0.05 (effect size = 0.076)	Women raised in a non-fluoridated area had significantly higher BMD of volumetric lumbar 3
	Saskatoon (Fluoridated)		33		0.227	Body weight	-		
Bone mineral density: Anterior-posterior lumbar spine									
Arnold *et al.* (1997)		g/cm^2, DXA in array mode		ANCOVA	Mean BMD		-		
	Regina (Non-fluoridated)		24		0.975	Body weight	-	p<0.05 (effect size = 0.119)	Women raised in non-fluoridated area had significantly higher density of anterior-posterior lumbar spine

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Saskatoon (Fluoridated)		33		1.039	Body weight	-		
Bone mineral density: Radius distal									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^^{2}, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure (reference)		1,243		0.353	Age		p=0.44	No association between fluoride exposure and density of distal radius
	1-10 years fluoride exposure		437		0.352	Age			
	11-20 years fluoride exposure		198		0.357	Age			
	>20 years fluoride exposure		192		0.363	Age			
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^^{3}, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.371	age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin	0.367–0.374	p=0.002	Women with continuous exposure had significantly lower density of the distal radius

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status			
	Continuous exposure for 20 years				0.364		0.361–0.367		
Bone mineral density: Radius proximal									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.634	Age		p=0.05 (1-10 years sig lower than 0 years)	Density of proximal radius significantly lower among those exposed for 1–10 years than those with 0 years exposure
	1-10 years fluoride exposure		437		0.622	Age			
	11-20 years fluoride exposure		198		0.635	Age			
	>20 years fluoride exposure		192		0.637	Age			
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^^{3}, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	No exposure in 20 years (reference)				0.645	Age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.642–0.649	p<0.001	Women with continuous exposure had significantly lower density of the proximal radius
	Continuous exposure for 20 years				0.636		0.633–0.639		
Bone mineral density: Radius calcaneus									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.404	Age		p=0.20	No association between fluoride exposure and density of radius calcaneus
	1-10 years fluoride exposure		437		0.395	Age			

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	11-20 years fluoride exposure		198		0.4	Age			
	>20 years fluoride exposure		192		0.405	Age			
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^3, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.408	Age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.405–0.412	p>0.05	No significant difference in density of the radius calcaneus between women with continuous exposure and those with no exposure
	Continuous exposure for 20 years				0.413		0.410–0.416		
Bone mineral density: Total hip									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^^{2}, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	0 years fluoride exposure		1,243		0.753	Age		p=0.24	No association between fluoride exposure and density of total hip
	1-10 years fluoride exposure		437		0.743	Age			
	11-20 years fluoride exposure		198		0.745	Age			
	>20 years fluoride exposure		192		0.766	Age			
Bone mineral density: Femoral head									
Chachra *et al.* (2010)		mg/cm^^{3}, micrometry		T-test	Mean BMD		Standard error of mean		
	Fluoridated region (Toronto)		53		0.9	-	0.04	p<0.05	Significantly higher BMD observed in samples from fluoridated region
	Non-fluoridated region (Montreal)		39		0.75	-	0.05		
Bone mineral density: Femoral neck									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.64	Age		p=0.32	No association between fluoride exposure and

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									density of femoral neck
	1-10 years fluoride exposure		437		0.638	Age			
	11-20 years fluoride exposure		198		0.642	Age			
	>20 years fluoride exposure		192		0.658	Age			
Lehmann *et al.* (1998)		g/cm^^{3}, DXA		MANOVA/ MANCOVA	Mean BMD		SD		No significant differences in adjusted BMD between regions
	Men fluoridated		41		0.876	Age and BMI	0.12	p=0.008 (NS, Bonferroni)	
	Men non-fluoridated		98		0.82	Age and BMI	0.101		
	Women fluoridated		201		0.809	Age and weight	0.102	p=0.65	
	Women non-fluoridated		215		0.814	Age and weight	0.1	p=0.008	
Phipps *et al.* (2000) (Pittsburgh)		g/cm^^{3}, method not specified		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.647	Age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen	0.643–0.651	p<0.001	Women with continuous exposure had significantly higher density of the femoral neck

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status			
	Continuous exposure for 20 years				0.664		0.661–0.668		
Kröger *et al.* (1994)		g/cm^3, DXA		ANCOVA with adjustment of confounding factors	Mean BMD		-		
	All women - fluoride		969		0.94	Age, weight, menopausal status, calcium intake, physical activity class, number of deliveries, alcohol consumption, and oestrogen use	-	p=0.004	Significantly higher density of the femoral neck observed for postmenopausal women
	All women - non-fluoride		2,253		0.93		-		
	Premenopausal women - fluoride		281†		0.963		-	p>0.05	
	Premenopausal women - non-fluoride		701†		0.953		-		
	Postmenopausal women - fluoride		688†		0.929		-	p=0.036	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Postmenopausal women - non-fluoride		1,552†		0.919		-		
Bone mineral density: Ward's triangle									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.426	Age		p=0.42	No association between fluoride exposure and density of Ward's triangle
	1-10 years fluoride exposure		437		0.421	Age			
	11-20 years fluoride exposure		198		0.428	Age			
	>20 years fluoride exposure		192		0.438	Age			
Lehmann *et al.* (1998)		g/cm^3, DXA		MANOVA/ MANCOVA	Mean BMD		SD		
	Men fluoridated		41		0.725	Age and BMI	0.129	p=0.002	Density of Ward's triangle significantly higher for men in fluoridated area; no effect among women
	Men non-fluoridated		98		0.655	Age and BMI	0.111		
	Women fluoridated		201		0.659	Age and weight	0.109	p=0.8	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Women non-fluoridated		215		0.662	Age and weight	0.11		
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^^{3}, method not specified		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.429	Age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.424–0.434	p=0.002	Women with continuous exposure had significantly higher density of Ward's triangle
	Continuous exposure for 20 years				0.436		0.436–0.443		
Bone mineral density: Trochanter									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.555	Age		p=0.15	No association between fluoride exposure and

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									density of the trochanter
	1-10 years fluoride exposure		437		0.548	Age			
	11-20 years fluoride exposure		198		0.547	Age			
	>20 years fluoride exposure		192		0.568	Age			
Lehmann *et al.* (1998)		g/cm^3, DXA		MANOVA/ MANCOVA	Mean BMD		SD		No significant differences in adjusted BMD between regions
	Men fluoridated		41		0.771	Age and BMI	0.105	p=0.01 (NS, Bonferroni)	
	Men non-fluoridated		98		0.724	Age and BMI	0.089		
	Women fluoridated		201		0.685	Age and weight	0.8	p=0.61	
	Women non-fluoridated		215		0.69	Age and weight	0.088		
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^^{3}, method not specified		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.558	Age, weight, education, knee/grip strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current	0.554–0.562	p<0.001	Women with continuous exposure had significantly higher density of the trochanter

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status			
	Continuous exposure for 20 years				0.572		0.568–0.575		
Bone mineral density: Intertrochanter									
Cauley *et al.* (1995) (Pittsburgh) ●		g/cm^2, single photon absorptiometry		ANOVA/ ANCOVA	Mean BMD				
	0 years fluoride exposure		1,243		0.883	Age		p=0.40	No association between fluoride exposure and density of the intertrochanter
	1-10 years fluoride exposure		437		0.873	Age			
	11-20 years fluoride exposure		198		0.878	Age			
	>20 years fluoride exposure		192		0.899	Age			
Phipps *et al.* (2000) (Pittsburgh) ●		g/cm^3, method not specified		ANOVA/ ANCOVA	Mean BMD		95% CI		
	No exposure in 20 years (reference)				0.892	Age, weight, education, knee/grip strength,	0.887–0.898	p>0.05	No significant difference in density of the

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status			intertrochanter between women with continuous and no exposure
	Continuous exposure for 20 years				0.895		0.889–0.900		
Bone mineral density: Proximal femur									
Arnold *et al.* (1997)		g/cm^^{2}, DXA in array mode		ANCOVA	Mean BMD		-		
	Regina (Non-fluoridated)		24		0.927	Body weight	-	p>0.05	No difference in density of proximal femur between women raised in fluoridated and non-fluoridated areas
	Saskatoon (Fluoridated)		33		0.961	Body weight	-		
Bone stiffness: Non-dominant foot									
O'Sullivan and		Quantitative ultrasound		Probit models estimated using	Marginal effect of prevalence		-		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
O'Connell (2014)				maximum likelihood estimation for having normal BMD, designated as bone stiffness > 86%	of fluoridated water on normal bone density				
	All participants		-		-0.033	Body mass index, whether one ever lived outside the Republic of Ireland, whether or not the respondent exercises at least 1 day/week, ever/currently smoke(d), self-report of growing up in a rural area, gender, coverage by private medical health insurance or government means–tested free medical care, age, residing in a non-completely urbanized electoral district (more than 1% of local labour force engaged in agriculture), the value of the	-	p>0.05	No association between prevalence of households with fluoridated water and probability of normal bone density

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						respondent's home, highest level of education completed, self-reported poor health when aged 14, self-report of family finances when aged 14 and local authority of residence (the +3500 electoral districts are subsets of the 34 local authorities).			
	Non-fully urbanised subsample (sensitivity analysis)		-		-0.027	Models repeated for sensitivity analysis; unclear if all potential confounding variables were included	-	p>0.05	
	Women only (sensitivity analysis)		-		-0.022	Models repeated for sensitivity analysis; unclear if all potential confounding variables were included	-	p>0.05	
	Men only (sensitivity analysis)		-		-0.037	Models repeated for sensitivity analysis; unclear if all potential confounding	-	p>0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						variables were included			
	Under 55s (sensitivity analysis)		-		-0.094	Models repeated for sensitivity analysis; unclear if all potential confounding variables were included	-	p>0.05	
Osteoporosis									
Lee *et al.* (2020)		Osteoporosis incidence data from National Health Insurance Service		Bayesian spatio-temporal regression analysis to calculate posterior relative risk	Relative risk		95% Credible Interval		
	Total sample		-		0.94	Space, time	0.87–1.02	p>0.05	Relative risks increased over time but did not increase in CWF area compared to non-CWF areas
	Male		-		0.86	Space, time	0.76–0.97	p>0.05	
	Female		-		0.95	Space, time	0.87–1.03	p>0.05	

Note: † denotes figures that were not extracted directly from papers but calculated by the review team based on provided information.

Table 78 Findings for studies examining fractures

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
All fractures									
Kröger *et al.* (1994)		Self-reported fracture since the age of 15 years		Student's two-tailed unpaired t-test	% participants reporting fracture		-		
	All women - fluoride		148		15.4	-	-	p>0.05	No difference in fracture incidence between fluoridated and non-fluoridated regions
	All women - non-fluoride		302		13.4	-	-		
Osteoporotic fracture									
Cauley *et al.* (1995) (Pittsburgh) ●		Non-spine fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Proportional hazard regression models, zero years exposure as reference group	Adjusted relative risk		95% CI		
	0 years fluoride exposure (reference)		178		1	Age, BMI, total calcium intake, history of osteoporosis, surgical menopause, history of falls in past year, drinks per week, education, current oestrogen use,	-	p=0.99	Women exposed for >20 years had about a 25% lower osteoporotic fracture risk, but CIs were wide and included 1.0.

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						current thiazide diuretic use, ever used bottled water			
	1–10 years fluoride exposure		69		1.04		0.77–1.38		
	11–20 years fluoride exposure		24		0.96		0.67–1.54		
	>20 years fluoride exposure		22		0.74		0.46–1.19		
	Any fluoride exposure		115				-		
Incidental spinal/vertebral fracture									
Cauley *et al.* (1995) (Pittsburgh) ●		Defined as 20% reduction in the vertebral height of the anterior, middle, or posterior dimension of a vertebral body and at least a 4 mm decrease in the vertebral height of a dimension, detected by repeat lateral and lumbar and thoracic vertebral film		Proportional hazard regression models, zero years exposure as reference group	Adjusted relative risk		95% CI		
	0 years fluoride exposure (reference)		58		1	Age, BMI, total calcium intake, history of osteoporosis, surgical menopause, history of falls in past year, drinks per week, education, current oestrogen use, current thiazide	-	-	No association between fluoride exposure and risk of incident vertebral fracture

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						diuretic use, ever used bottled water			
	1–10 years fluoride exposure		15		1.02		0.55–1.88		
	11–20 years fluoride exposure		5		0.58		0.21–1.60		
	>20 years fluoride exposure		4		1.63		0.57–4.67		
	Any fluoride exposure		24		-		-		
Phipps *et al.* (2000) (Pittsburgh) ●		Spine fractures, detected by lateral radiographs of thoracic and lumbar spine		Relative risk (continuous exposure using no exposure as reference group)	Relative risk		95% CI		
	No exposure in 20 years (reference)				1	Age, weight, education, muscle strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.55–0.97	p=0.033	Women with continuous exposure to fluoride had 27% lower risk of verbal fracture compared with those with no fluoride exposure
	Continuous exposure for 20 years				0.73				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Non-spine/non-vertebral fracture									
Cauley *et al.* (1995) (Pittsburgh) ●		Non-spine fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Proportional hazard regression models, zero years exposure as reference group	Adjusted relative risk		95% CI		
	0 years fluoride exposure (reference)		221		1	Age, BMI, total calcium intake, history of osteoporosis, surgical menopause, history of falls in past year, drinks per week, education, current oestrogen use, current thiazide diuretic use, ever used bottled water	-	p=0.99	Women exposed for >20 years had about a 25% lower risk of non-spinal fracture, but CIs were wide and included 1.0
	1–10 years fluoride exposure		89		1.1		0.85–1.42		
	11–20 years fluoride exposure		32		1.04		0.71–1.52		
	>20 years fluoride exposure		27		0.73		0.48–1.12		
	Any fluoride exposure		148		-		-		
Phipps *et al.* (2000) (Pittsburgh) ●		Non-spine fractures, self-reported, excluding fractures due to major		Relative risk, continuous exposure using no exposure as	Relative risk		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		trauma (e.g. motor vehicle accident)		reference group					
	No exposure in 20 years (reference)				1	Age, weight, education, muscle strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.83–1.10	p>0.05	No significant difference between women with continuous compared with no fluoride exposure
	Continuous exposure for 20 years				0.96				
Hip fracture									
Danielson *et al.* (1992)		Incidence of discharge for hip fracture among those aged 65 years or over		Age-adjusted relative risks using non-fluoridated community as referent	Risk ratio		95% CI		Small but statistically significant higher incidence of hip fractures in the fluoridated area for both men and women
	Female fluoridated		85		1.27	Age	1.08–1.46	-	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Female non-fluoridated		130		1				
	Male fluoridated		19		1.41		1.00–1.81	-	
	Male non-fluoridated		32		1				
Jacobsen *et al.* (1992)		Incidence of discharge for hip fracture among the white population aged 65 years or over		Age-adjusted rates calculated, Poisson regression models to estimate relative risk of hip fracture in fluoridated compared with non-fluoridated counties controlling for age	Risk ratio				Small positive ecological association between fluoridation of public water supplies and incidence of hip fracture among the aged population, although the authors cite numerous methodological concerns regarding accurate measurement and interpret their findings very cautiously
	Female fluoridated Female non-fluoridated				1.08	Age	1.06–1.1		
	Male fluoridated Male non-fluoridated				1.17	Age	1.13–1.22		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Suarez-Almazor *et al.* (1993)		Cases of hip fracture (discharge diagnosis) in individuals aged 45 years or over			Rate Ratio	Age, sex	95% CI		Generally no differences observed between the two regions - small difference for men total and men over 65, but authors judge this unlikely to be meaningful.
	Men 45-64 Edmonton (fluoridated)		-		1.07		0.87–1.32		
	Men 45-64 Calgary (non-fluoridated)		-						
	Men 65+ Edmonton (fluoridated)		-		1.13		1.00–1.27		
	Men 65+ Calgary (non-fluoridated)		-						
	Total men 45+ Edmonton (fluoridated)		-		1.12		1.01–1.24		
	Total men 45+ Calgary (non-fluoridated)		-						
	Women 45-64 Edmonton (fluoridated)		-		0.85		0.70–1.03		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Women 45-64 Calgary (non-fluoridated)		-						
	Women 65+ Edmonton (fluoridated)		-		0.96		0.90–1.03		
	Women 65+ Calgary (non-fluoridated)		-						
	Total Women 45+ Edmonton (fluoridated)		-		0.95		0.89–1.01		
	Total Women 45+ Calgary (non-fluoridated)		-						
	Both sexes 45-64 Edmonton (fluoridated)		-		0.94		0.82–1.08		
	Both sexes 45-64 Calgary (non-fluoridated)		-						
	Both sexes 65+ Edmonton (fluoridated)		-		1		0.94–1.06		
	Both sexes 65+ Calgary (non-fluoridated)		-						
	Total Both sexes 45+ Edmonton (fluoridated)		-		1		0.95–1.06		
	Total Both sexes 45+ Calgary (non-fluoridated)		-		1.07				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Jacobsen *et al.* (1993)		Incidence of hip fractures (i.e. proximal femur fracture) via Rochester Epidemiology Project and Mayo Clinic master index		Poisson regression analysis	Risk ratio		95% CI		No change in risk of hip fracture associated with fluoridation
	Both sexes–1950-1959 (Pre-fluoridation) Both sexes–1960-1969 (post-fluoridation)	751 (total pre and post)			0.63	Age, sex, and secular trend (controlled for in regression)	0.46–0.86		
	Women–1950-1959 (pre-fluoridation) Women–1960-1969 (post-fluoridation)	268 (total pre and post)			0.6	Age, sex, and secular trend (controlled for in regression)	0.42–0.85		
	Men–1950-1959 (pre-fluoridation)	383 (total pre and post)			0.78	Age, sex, and secular trend (controlled for in regression)	0.37–1.66		
	Men–1960-1969 (post-fluoridation)				0.63	Age, sex, and secular trend (controlled for in regression)	0.46–0.86		
Cauley *et al.* (1995) (Pittsburgh) ●		Hip fractures, self-reported and confirmed by review of copies of radiographs, excluding fractures due to major trauma (e.g. motor vehicle accident)		Proportional hazard regression models, zero years exposure as reference group	Adjusted relative risk		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	0 years fluoride exposure (reference)		27		1	Age, BMI, total calcium intake, history of osteoporosis, surgical menopause, history of falls in past year, drinks per week, education, current oestrogen use, current thiazide diuretic use, ever used bottled water	-	-	The relative risk of hip fracture tended to decrease with increasing duration of exposure to fluoride, but the CIs were wick and none of the relative risks were statistically significant.
	1–10 years fluoride exposure		10		0.89		0.42–1.92		
	11–20 years fluoride exposure		2		0.58		0.14–2.48		
	>20 years fluoride exposure		2		0.44		0.10–1.86		
	Any fluoride exposure				-		-		
Lehmann *et al.* (1998)		Incidence of hip fracture admissions to local hospitals			Odds ratio		95% CI		No difference in fracture incidence for those aged 35–59 years. No difference in fracture incidence between fluoridated and non-fluoridated areas for those

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									aged 65 years or over, except for women aged over 85 years, for whom fracture incidence was significantly lower in fluoridated areas.
	Men fluoridated 35-59 years Men non-fluoridated 35-59 years				1.01	Age			
	Women fluoridated 35-59 years Women non-fluoridated 35-59 years				0.97				
	Men fluoridated 60-64 Men non-fluoridated 60-64				2.14		0.89–5.20	p=0.09	
	Men fluoridated 65-69 Men non-fluoridated 65-69				0.55		0.22–1.39	p=0.19	
	Men fluoridated 70-74 Men non-fluoridated 70-74				0.78		0.27–1.39	p=0.64	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Men fluoridated 75-79 Men non-fluoridated 75-79				1.05		0.65–1.69	p=0.84	
	Men fluoridated 80-84 Men non-fluoridated 80-84				1.02		0.67–1.55	p=0.92	
	Men fluoridated >85 Men non-fluoridated >85				1.92		1.07–3.45	p=0.02 (NS, Bonferroni)	
	Women fluoridated 60-64 Women non-fluoridated 60-64				0.9		0.51–1.58	p=0.71	
	Women fluoridated 65-69 Women non-fluoridated 65-69				1.56		1.00–2.44	p=0.05 (NS, Bonferroni)	
	Women fluoridated 70-74 Women non-fluoridated 70-74				1.09		0.76–1.57	p=0.64	
	Women fluoridated 75-79 Women non-fluoridated 75-79				1.38		1.06–1.80	p=0.01 (NS, Bonferroni)	
	Women fluoridated 80-84 Women non-fluoridated 80-84				1.2		0.95–1.52	p=0.13	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Women fluoridated >85 Women non-fluoridated >85				1.41		1.10–1.81	p=0.006 (Significant, Bonferroni)	
Phipps *et al.* (2000) (Pittsburgh) ●		Hip fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Relative risk, continuous exposure using no exposure as reference group	Relative risk		95% CI		
	No exposure in 20 years (reference) Continuous exposure for 20 years				1 0.69	Age, weight, education, muscle strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.50–0.96	p=0.028	Women with continuous exposure to fluoride had 31% lower risk of hip fracture compared to those with no fluoride exposure
Young *et al.* (2015)		Number of hip fracture in-patient consultant episodes per lower super output area level in England recorded in		Incidence rate ratio	Incident rate ratio		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		hospital episode statistics between April 2007 and March 2013							
	All ages - fluoridated areas		45,219		0.7	Age, gender, deprivation, ethnicity	-1.0–2.4	p>0.05	No difference in rate of hip fractures between fluoridated and non-fluoridated areas
	All ages - non-fluoridated areas		303,848						
Lee *et al.* (2020)		Incidence of hip fracture, data gathered from National Health Insurance Scheme		Bayesian spatio-temporal regression analysis to calculate posterior relative risk	Risk ratio		95% Credible Interval		
	Total sample		-		0.95	Space, time	0.87–1.05	p>0.05	Relative risks increased over time but did not increase in CWF area compared to non-CWF areas
	Male		-		0.88	Space, time	0.75–1.01	p>0.05	
	Female		-		0.99	Space, time	0.89–1.09	p>0.05	
Wrist fracture									
Kröger *et al.* (1994)		Wrist fractures, self-reported, excluding fractures due to major		Student's two-tailed unpaired t-test	% participants ' reporting fracture		-		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		trauma (e.g. motor vehicle accident							
	All women - fluoride		43		4.4	-	-	p>0.05	No difference in fracture incidence between fluoride and non-fluoride groups
	All women - non-fluoride		70		3.1	-	-		
Cauley *et al.* (1995) (Pittsburgh) ●		Wrist fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Proportional hazard regression models, zero years exposure as reference group	Adjusted relative risk		95% CI		
	0 years fluoride exposure (reference)		44		1	Age, BMI, total calcium intake, history of osteoporosis, surgical menopause, history of falls in past year, drinks per week, education, current oestrogen use, current thiazide diuretic use, ever used bottled water	-	-	No association between fluoride exposure and risk of wrist fracture
	1–10 years fluoride exposure		19		1.17		0.67–2.05		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	11–20 years fluoride exposure		6		0.81		0.32–2.04		
	>20 years fluoride exposure		6		0.95		0.40–2.25		
	Any fluoride exposure		31		-		-		
Phipps *et al.* (2000) (Pittsburgh) ●		Wrist fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Relative risk, continuous exposure using no exposure as reference group	Relative risk		95% CI		
	No exposure in 20 years (reference)				1	age, weight, education, muscle strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	1.00–1.71	p>0.05	No significant difference between women with continuous compared with no fluoride exposure
	Continuous exposure for 20 years				1.32				
Ankle fracture									
Kröger *et al.* (1994)		Self-reported ankle fracture since the age of 15 years		Student's two-tailed unpaired t-test	% participants		-		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
					reporting fracture				
	All women - fluoride		24		2.5	-	-	p>0.05	No difference in fracture incidence between fluoride and non-fluoride groups
	All women - non-fluoride		49		2.2	-	-		
Non-wrist/non-ankle fracture									
Kröger *et al.* (1994)		Self-reported non-wrist/non-ankle fracture since age 15		Student's two-tailed unpaired t-test	% participants reporting fracture		-		
	All women - fluoride		81		8.4	-	-	p>0.05	No difference in fracture incidence between fluoride and non-fluoride groups
	All women - non-fluoride		183		8.1	-	-		
Humerus fracture									
Phipps *et al.* (2000) (Pittsburgh) ●		Humerus fractures, self-reported, excluding fractures due to major trauma (e.g. motor vehicle accident)		Relative risk, continuous exposure using no exposure as reference group	Relative risk		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	No exposure in 20 years (reference)				1	Age, weight, education, muscle strength, surgical menopause, calcium intake, drinks/week, current oestrogen use, current thiazide use, non-insulin dependent diabetes, current thyroid hormone use, walking for exercise, and smoking status	0.58–1.23	p>0.05	No significant difference between women with continuous compared with no fluoride exposure
	Continuous exposure for 20 years				0.85				

Table 79 Findings for studies examining neuropsychological outcomes

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Broadbent *et al.* (2015)									
FSIQ at the ages of 7-13 years (measured at 4 ages and averaged)		FSIQ, Wechsler Intelligence Scale for Children Revised (standardised to mean: 100; SD: 15)		General linear models, adjusting for confounders	Parameter estimate		95% CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Resident in CWF area (age 5 years)		891		-0.14	Sex, socioeconomic status in childhood, low birth weight, and breastfeeding	-3.49–3.2	p=0.93	No association between preschool fluoride exposure on IQ in childhood
	Never lived in CWF area (age 5 years)		99						
FSIQ at age 38 years		FSIQ, Wechsler Adult Intelligence Scale, Fourth Edition (standardised to mean: 100; SD ±15)		General linear models, adjusting for confounders	Parameter estimate		95% CI		
	Resident in CWF area (age 5 years)		847		3	Sex, socioeconomic status in childhood, low birth weight, breastfeeding, educational achievements	0.02–5.98	p=0.05	No association between preschool fluoride exposure on IQ in adulthood
	Never lived in CWF area (age 5 years)		93						
Verbal comprehension index at the age of 38 years		Verbal Comprehension Index, Wechsler Adult Intelligence Scale Revised (standardised to		General linear models	Mean		Standard deviation		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		mean: 100; SD: ±15)							
		Resident in CWF area (age 5 years)		891		100.1	-	14.9	p=0.51
		Never lived in CWF area (age 5 years)		99		98.9	-	15.5	
Perceptual reasoning index at the age of 38 years		Perceptual Reasoning Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)		General linear models	Mean		Standard deviation		
	Resident in CWF area (age 5 years)		891		100.2	-	15.1	p=0.18	No significant difference
	Never lived in CWF area (age 5 years)		99		98	-	14.1		
Working memory index at the age of 38 years		Working Memory Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)		General linear models	Mean		Standard deviation		
	Resident in CWF area (age 5 years)		891		100.3	-	15	p=0.12	No significant difference

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Never lived in CWF area (age 5 years)		99		97.7	-	15		
Processing speed index at age 38		Processing Speed Index, Wechsler Adult Intelligence Scale Revised (standardised to mean: 100; SD: ±15)		General linear models	Mean		Standard deviation		
	Resident in CWF area (age 5 years)		891		100.1	-	15.1	p=0.47	No significant difference
	Never lived in CWF area (age 5 years)		99		98.9	-	14.3		
Green *et al.* (2019) (MIREC) 🔴									
FSIQ at the ages of 3–4 years		FSIQ, Wechsler Preschool and Primary Scale of Intelligence, Third Edition, using USA population-based normative data (mean: 100; SD: 15).			Adjusted mean change associated with 1mg change in predictor variable		95%CI		
	Maternal urinary fluoride sample		512	Linear regression models; predictor variable	-1.95	City, HOME score, maternal education, race/ethnicity, and	-5.19–1.28	-	No association between maternal fluoride and IQ

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				maternal urinary fluoride, specific gravity-adjusted		including child sex interaction.			score total sample
	Maternal urinary fluoride sample Boys		-	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.49	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.38–-0.60	p<0.05	1mg increase in maternal urinary fluoride associated with a 4.49 lower FSIQ score in boys
	Maternal urinary fluoride sample Girls		-	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	2.4	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-2.53–7.33	p>0.05	No association between maternal fluoride and IQ score in girls
	Fluoride intake sample		400	Linear regression models; predictor variable fluoride intake from tap water and tea/coffee	-3.66	City, HOME score, maternal education, race/ethnicity, child sex, and prenatal secondhand smoke exposure	-7.16–-0.15	p<0.05	1mg increase in fluoride intake associated with a 3.66 lower FSIQ score among boys and girls

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Water fluoride concentration sample		420	Linear regression models; predictor variable water fluoride concentration in mg/L	-5.29	HOME score, maternal education, race, child sex, and prenatal second-hand smoke exposure; because city was strongly multi-collinear with water fluoride concentration (VIF >20), it was excluded from the model	-10.39–-0.19	p<0.05	1mg increase in fluoride concentration associated with a 5.29 lower FSIQ score among boys and girls
	Maternal urinary fluoride sample Boys (sensitivity analyses)		504	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.61	Lead plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.50–-0.71	p<0.05	
	Maternal urinary fluoride sample Boys (sensitivity analyses)		456	Linear regression models; predictor variable maternal urinary fluoride,	-5.13	Mercury plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-9.16–-1.10	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				specific gravity-adjusted					
	Maternal urinary fluoride sample Boys (sensitivity analyses)		503	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.57	Perfluorooctanoic acid plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.21–-0.50	p<0.05	
	Maternal urinary fluoride sample Boys (sensitivity analyses)		512	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.44	Arsenic plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.35–-0.54	p<0.05	
	Maternal urinary fluoride sample Boys (sensitivity analyses)		502	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.55	Manganese plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.42–-0.69	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Maternal urinary fluoride sample Boys (sensitivity analyses)		512	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.18	Secondhand smoke exposure plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.06–-0.30	p<0.05	
	Maternal urinary fluoride sample Boys (sensitivity analyses)		407	Linear regression models; predictor variable maternal urinary fluoride, creatinine adjusted	-4.96	Lead plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-8.56–-1.36	p=0.007	
	Maternal urinary fluoride sample Boys (sensitivity analyses)		369	Linear regression models; predictor variable water fluoride concentration in mg/L	-6.25	Lead plus city, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-11.56–-0.94	p=0.02	
	Maternal urinary fluoride sample Boys		510	Linear regression models; predictor variable	-4.11	City, HOME score, maternal education, race/ethnicity, and	-7.89–-0.33	p=0.03	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	(sensitivity analyses)			maternal urinary fluoride, specific gravity-adjusted		including child sex interaction.			
Verbal IQ at the ages of 3–4 years		Verbal IQ, Weschler Preschool and Primary Scale of Intelligence, Third Edition, Verbal IQ			Adjusted mean change associated with 1mg change in predictor variable		95%CI		No association between change in maternal fluoride or fluoride intake or water fluoride concentration and verbal IQ
	Maternal urinary fluoride sample		509	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-1.6	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-4.74, 1.55	p>0.05	
	Maternal urinary fluoride sample Boys			Linear regression models; predictor variable maternal urinary fluoride,	-2.85	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-6.65, 0.95	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				specific gravity-adjusted					
	Maternal urinary fluoride sample Girls			Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	0.55	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-4.28, 5.37	p>0.05	
	Fluoride intake sample		395	Linear regression models; predictor variable fluoride intake from tap water and tea/coffee	-3.08	City, HOME score, maternal education, race/ethnicity, child sex, and prenatal secondhand smoke exposure	-6.40, 0.25	p>0.05	
	Water fluoride concentration sample		420	Linear regression models; predictor variable water fluoride concentration in mg/L	-1.6	City, HOME score, maternal education, race/ethnicity, and including child sex interaction.	-4.74, 1.55	p>0.05	
Performance IQ at the ages of 3–4 years		Performance IQ, Weschler Preschool and Primary Scale of Intelligence,			Adjusted mean change associated with 1mg		95%CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		Third Edition, Performance IQ			change in predictor variable				
	Maternal urinary fluoride sample		507	Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-1.24	City, HOME score, maternal education, race and including child sex interaction	-4.88–2.4	p>0.05	
	Maternal urinary fluoride sample Boys			Linear regression models; predictor variable maternal urinary fluoride, specific gravity-adjusted	-4.63	City, HOME score, maternal education, race and including child sex interaction	-9.01–-0.25	p<0.05	Increase of 1mg/L maternal urinary fluoride SG associated with a 4.63 lower performance IQ score in boys, but no difference in girls.
	Maternal urinary fluoride sample Girls			Linear regression models; predictor variable maternal urinary fluoride,	4.51	City, HOME score, maternal education, race and including child sex interaction	-1.02–10.05	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				specific gravity-adjusted					
	Fluoride intake sample		399	Linear regression models; predictor variable fluoride intake from tap water and tea/coffee	-2.74	HOME score, maternal education, race, child sex, prenatal second-hand smoke exposure, and city	-6.82–1.34	p>0.05	No effect of change in fluoride intake.
	Water fluoride concentration sample		420	Linear regression models; predictor variable water fluoride concentration in mg/L	-13.79	HOME score, maternal education, race, child sex, and prenatal second-hand smoke exposure; because city was strongly multi-collinear with water fluoride concentration (VIF >20), it was excluded from the model	-18.82–-7.28	p>0.05	Increased water fluoridation associated with lower performance IQ score for total sample
Till *et al.* (2020) (MIREC) 🔴									
FSIQ at the ages of 3–4 years		Full Scale IQ score differences, Wechsler Preschool and Primary Scale of Intelligence-III (Wechsler, 2002)			Adjusted mean difference associated with 0.5mg change in in predictor variable		95%CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		using United States population-based normative data (mean=100, SD=15).							
	Formula-fed group		398 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	-4.4	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-8.34–-0.46	p<0.05	Higher water fluoridation associated with lower FSIQ scores for formula-fed group only, no difference observed for breastfed group. This difference was no longer significant when maternal urinary fluoride was controlled for.
	Breastfed group			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding	-1.34	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and	-5.04–2.38	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				status (breast or formula)		second-hand smoke status in the child's house (yes, no)			
	Formula-fed group		350 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-3.58	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-7.83–0.66	p>0.05	
	Breastfed group			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-1.69	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-5.66–2.27	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Formula-fed group		398	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	-2.69	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-7.38–2.01	p>0.05	
	Formula-fed group		350	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula, controlling for maternal urinary fluoride	-1.94	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-7.09–3.21	p>0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		396 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride	-3.14	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME	-6.99–0.71	p<0.05	Association of water fluoridation and FSIQ no longer significant after removal of two IQ outliers

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				concentration, by feeding status (breast or formula)		total score (continuous), and second-hand smoke status in the child's house (yes, no)			
	Breastfed group (sensitivity analysis without extreme IQ outliers)			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	-1.38	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-4.97–2.22	p>0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		349 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-2.82	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-7.00–1.35	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Breastfed group (sensitivity analysis without extreme IQ outliers)			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-1.69	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-5.58–2.19	p>0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		396	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	-1.12	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-6.17–2.93	p>0.05	
	Formula-fed group (sensitivity analysis without		349	Linear regression, outcome variable IQ score, predictor variable	-1.2	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous),	-6.24–3.85	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	extreme IQ outliers)			fluoride intake from formula, controlling for maternal urinary fluoride		child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)			
Verbal IQ at the ages of 3–4 years		Verbal Scale IQ score differences, Wechsler Preschool and Primary Scale of Intelligence-III (Wechsler, 2002) using United States population-based normative data (mean=100, SD=15).			Adjusted mean difference associated with 0.5mg change in in predictor variable		95%CI		
	Formula-fed group		397 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	0.89	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the	-2.87–4.65	p>0.05	Water fluoride concentration not associated with changes in verbal IQ for either formula or breastfed groups, remaining statistically non-significant when

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						child's house (yes, no)			controlling for maternal urinary fluoride
	Breastfed group			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	3.06	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-.049–6.61	p>0.05	
	Formula-fed group		349 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	2.6	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-1.98–7.16	p>0.05	
	Breastfed group			Linear regression, outcome	4.2	Maternal education (binary), maternal race	-0.06–8.45	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride		(binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)			
	Formula-fed group		397	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	3.08	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-1.40–7.55	p>0.05	
	Formula-fed group		349	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula, controlling for	3.05	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and	-1.89–7.98	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				maternal urinary fluoride		second-hand smoke status in the child's house (yes, no)			
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		395 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	2.07	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-1.60–5.74	p>0.05	Association remained non-significant after removal of outliers
	Breastfed group (sensitivity analysis without extreme IQ outliers)			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	3.03	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-0.41–6.46	p>0.05	
	Formula-fed group (sensitivity		348 (both groups)	Linear regression, outcome	1.65	Maternal education (binary), maternal race	-2.35–5.65	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	analysis without extreme IQ outliers)			variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride		(binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)			
	Breastfed group (sensitivity analysis without extreme IQ outliers)			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	2.42	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-1.31–6.14	p>0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		395	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	4.08	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score	-0.26–8.42	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						(continuous), and second-hand smoke status in the child's house (yes, no)			
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		348	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula, controlling for maternal urinary fluoride	3.77	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-1.06–8.60	p>0.05	
Performance IQ at the ages of 3–4 years		Performance IQ score differences, Wechsler Preschool and Primary Scale of Intelligence-III (Wechsler, 2002) using United States population-based normative data (mean=100, SD=15).			Adjusted mean difference associated with 0.5mg change in in predictor variable		95%CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Formula-fed group		393 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	-9.26	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-13.77–-4.76	p<0.05	Water fluoride concentration significantly associated with lower performance IQ in both formula and breastfed groups, remaining statistically significant when controlling for maternal urinary fluoride
	Breastfed group			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	-6.19	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-10.45–-1.94	p<0.05	
	Formula-fed group		345 (both groups)	Linear regression, outcome variable IQ	-7.93	Maternal education (binary), maternal race (binary), child's age	-12.84–-3.01	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride		at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)			
	Breastfed group			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-6.3	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-10.92–-1.68	p<0.05	
	Formula-fed group		393	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	-8.76	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and	-14.18–-3.34	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						second-hand smoke status in the child's house (yes, no)			
	Formula-fed group		345	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula, controlling for maternal urinary fluoride	-7.62	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-13.64–-1.6	p<0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		391 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)	-8.23	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-12.70–-3.77	p<0.05	Association remains significant after removal of IQ outliers
	Breastfed group (sensitivity			Linear regression, outcome	-6.22	Maternal education (binary), maternal race	-10.41–-2.04	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	analysis without extreme IQ outliers)			variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula)		(binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)			
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		344 (both groups)	Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding status (breast or formula), controlling for maternal urinary fluoride	-7.31	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-12.20–-2.43	p<0.05	
	Breastfed group (sensitivity analysis without extreme IQ outliers)			Linear regression, outcome variable IQ score, predictor variable water fluoride concentration, by feeding	-6.29	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and	-10.86–-1.72	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				status (breast or formula), controlling for maternal urinary fluoride		second-hand smoke status in the child's house (yes, no)			
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		391	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula	-7.88	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-14.18–-3.34	p<0.05	
	Formula-fed group (sensitivity analysis without extreme IQ outliers)		344	Linear regression, outcome variable IQ score, predictor variable fluoride intake from formula, controlling for maternal urinary fluoride	-7.01	Maternal education (binary), maternal race (binary), child's age at IQ testing (continuous), child's sex, HOME total score (continuous), and second-hand smoke status in the child's house (yes, no)	-12.98–-1.03	p<0.05	
Farmus *et al.* (2021) (MIREC) 🔴									

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
FSIQ at the ages of 3–4 years		Full Scale IQ, Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III; Canadian norms; Wechsler, 2002)		Generalised estimating equation (GEE); predictor variable standardised maternal urinary fluoride	B coefficient		95% CI		
	Males		291		-1.86	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.22 — -0.49	p<0.05	Association between prenatal exposure to fluoride and FSIQ for boys
	Females		305		-0.23	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.06 — 1.60	p>0.05	
	Total sample		596		-1.28	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.37 — -0.18	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				GEE; predictor variable standardised infant fluoride intake	B coefficient		95% CI		
	Males		218		-0.01	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.67 — 1.65	p>0.05	No association between exposure to fluoride during infancy and FSIQ
	Females		214		-0.72	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.34 — 0.89	p>0.05	
	Total sample		432		-0.38	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.53 — 0.78	p>0.05	
				GEE; predictor variable standardised	B coefficient		95% CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				child urinary fluoride					
	Males		211		0.07	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.66 — 1.80	p>0.05	No association between childhood exposure to fluoride and FSIQ
	Females		223		-0.41	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.07 — 1.24	p>0.05	
	Total sample		434		-0.18	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.38 — 1.02	p>0.05	
				GEE; interaction between fluoride exposure (maternal urinary	X2 (df)				

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				fluoride) and time					
	Total sample		434		4.36 (3)			p=0.23	Interaction between exposure timing and fluoride level was not significant
				GEE; predictor variable standardised maternal urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		287		-1.22	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.41 — -0.04	p>0.05	Following removal of influential dyads, the association between MUF and FSIQ among boys was weaker and no longer significant.
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if		302		-1	Maternal education, maternal race, total HOME score, age at urine sampling, and	-2.84 — 0.84	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	removal would change co-efficients of exposure variables by at least 0.4 SDs)					prenatal second-hand smoke			
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		589		-1.14	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.25 — -0.04	p>0.05	
				GEE; predictor variable standardised infant fluoride intake	B coefficient				
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-		287		0.1	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.55 — 1.75	p>0.05	No association between exposure to fluoride during infancy and FSIQ (consistent with main analysis)

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	efficients of exposure variables by at least 0.4 SDs)								
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		302		-1.58	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.17 — 0.01	p>0.05	
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		589		-0.76	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.89 — 0.38	p>0.05	
				GEE; predictor variable standardised	B coefficient				

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				child urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		287		0.4	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.14 — 1.95	p>0.05	No association between childhood exposure to fluoride and FSIQ (consistent with main analysis)
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		302		0	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.61 — 1.61	p>0.05	
	Total sample (Sensitivity analysis; excluding influential		589		0.18	Maternal education, maternal race, total HOME score, age at urine	-1.01 — 1.38	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)					sampling, and prenatal second-hand smoke			
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		589		-			p=0.08	Interaction between exposure timing and fluoride level was not significant (consistent with main analysis)
				GEE; effect of 0.5mg/L change in average					

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				maternal urinary fluoride					
	Males		291		-2.48	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.30 — -0.66	p<0.05	Association between prenatal exposure to fluoride and FSIQ for boys only, consistent with analysis of standardised MUF
	Females		305		-0.31	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.76 — 2.14	p>0.05	
	Total sample		596		-1.71	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.17 — -0.24	p>0.05	
				GEE; effect of 0.1mg/day change in infant fluoride intake	B coefficient		95% CI		

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Males		218		-0.01	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.25 — 1.24	p>0.05	No association between exposure to fluoride during infancy and verbal IQ, consistent with analysis of standardised IFI
	Females		214		-0.54	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.75 — 0.66	p>0.05	
	Total sample		432		-0.28	Maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.15 — 0.58	p>0.05	
				GEE; effect of 0.5mg/L change in child urinary fluoride	B coefficient		95% CI		
	Males		211		0.09	Maternal education, maternal race, total HOME score,	-2.10 — 2.28	p>0.05	No association between childhood exposure to

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						age at urine sampling, and prenatal second-hand smoke			fluoride and verbal IQ, consistent with analysis of standardised CUF
	Females		223		-0.52	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.62 — 1.58	p>0.05	
	Total sample		434		-0.23	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.75 — 1.29	p>0.05	
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				
	Total sample		434		4.36 (3)			p=0.23	Interaction between

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									exposure timing and fluoride level was not significant
Verbal IQ at the ages of 3–4 years		Verbal Scale IQ, Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III; Canadian norms; Wechsler, 2002)		GEE; predictor variable standardised maternal urinary fluoride					
	Males		291		-0.25	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.57 — 1.07	p>0.05	No association between prenatal exposure to fluoride and verbal IQ
	Females		305		0.87	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.91 — 2.64	p>0.05	
	Total sample		596		0.15	maternal education, maternal race, total HOME score, age at urine	-0.91 — 1.20	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						sampling, and prenatal second-hand smoke			
				GEE; predictor variable standardised infant fluoride intake					
	Males		218		1.22	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.39 — 2.83	p>0.05	No association between exposure to fluoride during infancy and verbal IQ
	Females		214		1.31	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.25 — 2.87	p>0.05	
	Total sample		432		1.27	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	0.15 — 2.39	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				GEE; predictor variable standardised child urinary fluoride					
	Males		211		1.61	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.06 — 3.29	p>0.05	No association between childhood exposure to fluoride and verbal IQ
	Females		223		0.63	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.98 — 2.23	p>0.05	
	Total sample		434		1.1	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.06 — 2.26	p>0.05	
				GEE; interaction between fluoride	X2 (df)				

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				exposure (maternal urinary fluoride) and time					
	Total sample		434		4.36 (3)			p=0.04	Interaction between exposure timing and fluoride level was significant
				GEE; predictor variable standardised maternal urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		288		0.25	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.11 — 1.61	p>0.05	No association between prenatal exposure to fluoride and verbal IQ (consistent with main analysis)
	Females (Sensitivity analysis; excluding		302		0.33	maternal education, maternal race, total HOME score,	-1.47 — 2.13	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)					age at urine sampling, and prenatal second-hand smoke			
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		590		0.28	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.80 — 1.36	p>0.05	
				GEE; predictor variable standardised infant fluoride intake					
	Males (Sensitivity analysis; excluding influential dyads with		288		1.35	maternal education, maternal race, total HOME score, age at urine sampling, and	-0.24 — 2.93	p>0.05	No association between exposure to fluoride during infancy and verbal IQ

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)					prenatal second-hand smoke			(consistent with main analysis)
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		302		0.64	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.91 — 2.19	p>0.05	
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		590		0.99	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.12 — 2.09	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				GEE; predictor variable standardised child urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		288		1.89	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	0.16 — 3.62	p>0.05	No association between childhood exposure to fluoride and verbal IQ (consistent with main analysis)
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		302		0.98	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.60 — 2.55	p>0.05	
	Total sample (Sensitivity		590		1.39	maternal education,	0.23 — 2.56	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)					maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke			
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		590	-				p=0.03	Interaction between exposure timing and fluoride level was significant (consistent with main analysis)

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				GEE; effect of 0.5mg/L change in average maternal urinary fluoride					
	Males		291		-0.34	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.10 — 1.43	p>0.05	No association between prenatal exposure to fluoride and verbal IQ, , consistent with analysis of standardised MUF
	Females		305		1.16	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.22 — 3.53	p>0.05	
	Total sample		596		0.2	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.22 — 1.61	p>0.05	
				GEE; effect of 0.1mg/day					

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				change in infant fluoride intake					
	Males		218		0.92	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.29 — 2.12	p>0.05	No association between exposure to fluoride during infancy and verbal IQ, , consistent with analysis of standardised IFI
	Females		214		0.98	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.19 — 2.15	p>0.05	
	Total sample		432		0.95	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	0.11 — 1.79	p>0.05	
				GEE; effect of 0.5mg/L change in child urinary fluoride					
	Males		211		2.05	maternal education,	-0.08 — 4.16	p>0.05	No association between

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke			childhood exposure to fluoride and verbal IQ, , consistent with analysis of standardised CUF
	Females		223		0.79	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-1.24 — 2.82	p>0.05	
	Total sample		434		1.39	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-0.08 — 2.86	p>0.05	
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Total sample		434		8.28(3)			p=0.04	Interaction between exposure timing and fluoride level was significant
Performance IQ at the ages of 3–4 years		Performance Scale IQ, Wechsler Preschool and Primary Scale of Intelligence-III (WPPSI-III; Canadian norms; Wechsler, 2002)		GEE; predictor variable standardised maternal urinary fluoride	B coefficient		95% CI		
	Males		291		-3.01	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.60 — -1.42	p<0.05	Association between prenatal exposure to fluoride and performance IQ for whole sample and for boys
	Females		305		-1.18	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.32 — 0.96	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Total sample		596		-2.36	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.63 — -1.08	p<0.05	
				GEE; predictor variable standardised infant fluoride intake					
	Males		218		-1.45	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.40 — 0.49	p>0.05	Association between exposure to fluoride during infancy and performance IQ for whole sample and for girls
	Females		214		-2.71	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.59 — -0.83	p<0.05	
	Total sample		432		-2.11	maternal education, maternal race,	-3.45 — -0.76	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						total HOME score, age at urine sampling, and prenatal second-hand smoke			
				GEE; predictor variable standardised child urinary fluoride					
	Males		211		-1.49	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.50 — 0.53	p>0.05	No association between childhood exposure to fluoride and performance IQ
	Females		223		-1.53	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.45 — 0.39	p>0.05	
	Total sample		434		-1.51	maternal education, maternal race, total HOME score, age at urine sampling, and	-2.90 — -0.12	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						prenatal second-hand smoke			
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				
	Total sample		434		18.78 (3)			p<0.001	Interaction between exposure timing and fluoride level was highly significant
				GEE; predictor variable standardised maternal urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure		288		-2.39	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.05 — -0.73	p<0.05	Association between prenatal exposure to fluoride and PIQ for whole sample and for boys (consistent with main analysis)

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	variables by at least 0.4 SDs)								
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		301		-2	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.19 — 0.20	p>0.05	
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		589		-2.24	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.56 — -0.92	p<0.05	
				GEE; predictor variable standardised infant fluoride intake					

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		288		-1.38	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.32 — 0.55	p>0.05	Association between exposure to fluoride during infancy and PIQ for whole sample and for girls (consistent with main analysis)
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		301		-3.59	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-5.48 — -1.70	p<0.05	
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if		589		-2.51	maternal education, maternal race, total HOME score, age at urine sampling, and	-3.86 — -1.16	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	removal would change co-efficients of exposure variables by at least 0.4 SDs)					prenatal second-hand smoke			
				GEE; predictor variable standardised child urinary fluoride					
	Males (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		288		-1.17	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.29 — 0.94	p>0.05	No association between childhood exposure to fluoride and PIQ (consistent with main analysis)
	Females (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-		301		-1.21	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.12 — 0.71	p>0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	efficients of exposure variables by at least 0.4 SDs)								
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if removal would change co-efficients of exposure variables by at least 0.4 SDs)		589		-1.19	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.61 — 0.23	p>0.05	
				GEE; interaction between fluoride exposure (maternal urinary fluoride) and time	X2 (df)				
	Total sample (Sensitivity analysis; excluding influential dyads with FSIQ <70 or if		589	-				p<0.0001	Interaction between exposure timing and fluoride level was highly significant

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	removal would change co-efficients of exposure variables by at least 0.4 SDs)			GEE; effect of 0.5mg/L change in average maternal urinary fluoride					(consistent with main analysis)
	Males		291		-4.02	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-6.15 — -1.89	p<0.05	Association between prenatal exposure to fluoride and performance IQ for boys and for whole sample
	Females		305		-1.58	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.43 — 1.28	p>0.05	
	Total sample		596		-3.15	maternal education, maternal race, total HOME score, age at urine sampling, and	-4.85 — -1.44	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						prenatal second-hand smoke			
				GEE; effect of 0.1mg/day change in infant fluoride intake					
	Males		218		-1.09	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.54 — 0.37	p>0.05	Association between exposure to fluoride during infancy and performance IQ for whole sample and for girls, consistent with analysis of standardised IFI
	Females		214		-2.03	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.43 — -0.63	p<0.05	
	Total sample		432		-1.58	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-2.59 — -0.57	p<0.05	

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				GEE; effect of 0.5mg/L change in child urinary fluoride					
	Males		211		-1.89	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.44 — 0.67	p>0.05	No association between childhood exposure to fluoride and performance IQ, consistent with analysis of standardised CUF
	Females		223		-1.94	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-4.37 — 0.50	p>0.05	
	Total sample		434		-1.91	maternal education, maternal race, total HOME score, age at urine sampling, and prenatal second-hand smoke	-3.68 — -0.15	p>0.05	
				GEE; interaction between					

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				fluoride exposure (maternal urinary fluoride) and time					
	Total sample		434		18.78(3)			p<0.001	Interaction between exposure timing and fluoride level was highly significant
Barberio *et al.* (2017a) (CHMS)									
Learning disability age 3-12 years		Parental- or self-reported diagnosis of a learning disability			Adjusted Odds ratio		95% CI		
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on creatinine-adjusted urinary fluoride	1.04	Age, sex, household income adequacy, highest attained education in the household	0.95 –1.15	p>0.05	Reported learning disability diagnosis was not significantly associated with creatinine-adjusted urinary fluoride, specific gravity-adjusted urinary fluoride or fluoride concentration of tap water in

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									adjusted or unadjusted models
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	1.01	Age, sex, household income adequacy, highest attained education in the household	0.99–1.02	p>0.05	
	Cycle 3, fluoride urine subsample		-	Logistic regression, outcome regressed on creatinine-adjusted urinary fluoride	1.03	Age, sex, household income adequacy, highest attained education in the household	0.86–1.23	p>0.05	
	Cycle 3, fluoride urine subsample		-	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	1.01	Age, sex, household income adequacy, highest attained education in the household	0.99–1.03	p>0.05	
	Cycle 3, fluoride tap water subsample		-	Logistic regression, outcome regressed on specific fluoride tap water	0.88	Age, sex, household income adequacy, highest attained education in the household	0.068–11.33	p>0.05	
	Pooled Cycles 2 and 3,		-	Logistic regression, outcome	1.04	Age, sex, household income adequacy, highest	0.98–1.10	p>0.05	Reported learning disability

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	fluoride urine subsamples			regressed on creatinine-adjusted urinary fluoride		attained education in the household			diagnosis was not significantly associated with creatinine-adjusted urinary fluoride, specific gravity-adjusted urinary fluoride or fluoride concentration of tap water in adjusted or unadjusted models
	Pooled Cycles 2 and 3, fluoride urine subsamples		-	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	1.01	Age, sex, household income adequacy, highest attained education in the household	1.00–1.02	p>0.05	
ADHD diagnosis in childhood/ adolescence		Parental- or self-reported diagnosis of ADHD, age 3-12 years			Odds ratio		95% CI		
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on creatinine-	1.01	Age, sex, household income adequacy, highest attained education in the household	0.85–1.21	p>0.05	Reported diagnosis of ADHD not associated with either urinary measure of

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				adjusted urinary fluoride					fluoride exposure
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	1.01	Age, sex, household income adequacy, highest attained education in the household	0.96–1.06	p>0.05	
ADD diagnosis age 3-12 years		Parental- or self-reported diagnosis of ADD (no hyperactivity), age 3-12 years							
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on creatinine-adjusted urinary fluoride	0.79	Age, sex, household income adequacy, highest attained education in the household	0.59–1.06	p>0.05	Reported diagnosis of ADD significantly associated with creatinine-adjusted urinary fluoride, such that those with higher creatinine-adjusted urinary fluoride had lower odds of reporting ADD; however, this association was reduced to

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									non-significance in the adjusted model.
	Cycle 2, fluoride urine subsample		-	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	0.98	Age, sex, household income adequacy, highest attained education in the household	0.94–1.03	p>0.05	
Riddell *et al.* (2019) (CHMS) 🟢									
ADHD diagnosis in childhood/ adolescence		Physician-made diagnosis of ADHD age 6-17 years			Odds ratio		95% CI		
	Urinary fluoride sample		1877	Logistic regression, outcome regressed on specific gravity-adjusted urinary fluoride	0.96	Child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	0.63–1.46	p>0.05	Urinary fluoride did not significantly predict ADHD diagnosis
	Fluoride in tap water sample		710	Logistic regression, outcome regressed on fluoride	6.1	Child's sex, age at interview, ethnicity (white or other), body mass index, highest level of	1.60–22.8	p<0.05	1mg/L increase in tap water fluoride was associated with a 6.1 (1.60–

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				concentration in tap water		parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level			22.8) times higher odds of ADHD diagnosis, no interaction with age or sex
	CWF status sample		1722	Logistic regression, outcome regressed on CWF status	1.21	Child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	1.03–1.42	p<0.05	Significant interaction between age and CWF status, such that for older youth (75th percentile), predicted odds of ADHD diagnosis was 2.84 (1.40–5.76) times higher among youth in fluoridated region than non-fluoridated region; no difference across regions in odds for youth in 25th percentile

Outcome	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	CWF status sample - 75th age percentile				2.84	Child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	1.40–5.76	p<0.05	
	CWF status sample - 25th age percentile				0.91	Child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	0.41–1.99	p>0.05	

Table 80 Findings for studies examining neuropsychological outcomes using regression analysis

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
Neuropsychological development at 1 year									

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
Ibarluzea *et al.* 2021		Standardised scores Bayley Mental Development Index (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
			316		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	2.67	-3.46–8.81	p>0.05	No significant association between maternal fluoride during pregnancy and scores on Bayley's Mental Development Index at 1 year, including adjustment for blood cord mercury levels
					Maternal urinary fluoride (first trimester); creatinine-adjusted	0.89	-4.55–6.32	p>0.05	
					Maternal urinary fluoride (third trimester); creatinine-adjusted	2.65	-2.14–7.45	p>0.05	

Neuropsychological development at 4 years - Verbal

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
Ibarluzea 2021		Standardised scores on McCarthy Scales of Children's Abilities (Verbal scale), adapted to Spanish population (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for age of the child at time of test, order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
	Boys		125		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	9.4	-1.78–20.57	p>0.05	1mg increase in maternal urinary fluoride across whole pregnancy associated with a 9.74-point (1.75 to 17.74) higher verbal score in boys when adjusted for blood cord mercury levels; no association for girls.
	Girls		123		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	-2.07	-10–5.87	p>0.05	
	Total sample		248		Maternal urinary fluoride (first trimester);	-1.5	-7.53–4.54	p>0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
					creatinine-adjusted				
	Boys		125		Maternal urinary fluoride (third trimester); creatinine-adjusted	9.74	1.75–17.74	p<0.05	
	Girls		123		Maternal urinary fluoride (third trimester); creatinine-adjusted	-0.74	-6.72–5.25	p>0.05	
Neuropsychological development at 4 years - Performance									
Ibarluzea 2021		Standardised scores on McCarthy Scales of Children's Abilities (Performance scale), adapted to Spanish population (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for age of the child at time of test, order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
	Total sample		248		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	4.41	-1.59–10.41	p>0.05	No significant association between maternal fluoride during pregnancy and performance scores at 4 years when adjusted

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
									for blood cord mercury levels
			248		Maternal urinary fluoride (first trimester); creatinine-adjusted	3.85	-1.62–9.33	p>0.05	
			248		Maternal urinary fluoride (third trimester); creatinine-adjusted	2.33	-2.15–6.82	p>0.05	
Neuropsychological development at 4 years - Numeric									
Ibarluzea 2021		Standardised scores on McCarthy Scales of Children's Abilities (Numeric scale), adapted to Spanish population (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for age of the child at time of test, order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
	Total sample		248		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	5.28	-0.54–11.1	p>0.05	No significant association between maternal fluoride during pregnancy and numeric scores at 4 years when adjusted

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
									for blood cord mercury levels
			248		Maternal urinary fluoride (first trimester); creatinine-adjusted	3.38	-1.96–8.71	p>0.05	
			248		Maternal urinary fluoride (third trimester); creatinine-adjusted	3.47	-0.88–7.82	p>0.05	
Neuropsychological development at 4 years - Memory									
Ibarluzea 2021		Standardised scores on McCarthy Scales of Children's Abilities (Memory scale), adapted to Spanish population (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for age of the child at time of test, order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
	Total sample		248		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	0.8	-5.3–6.9	p>0.05	No significant association between maternal fluoride during pregnancy and memory scores at 4 years when adjusted

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
									for blood cord mercury levels
			248		Maternal urinary fluoride (first trimester); creatinine-adjusted	-0.52	-6.06–5.02	p>0.05	
			248		Maternal urinary fluoride (third trimester); creatinine-adjusted	1.15	-3.4–5.69	p>0.05	
Neuropsychological development at 4 years - General cognitive									
Ibarluzea 2021		Standardised scores on McCarthy Scales of Children's Abilities (General cognitive scale), adapted to Spanish population (standardised to mean 100 and SD 15)		Multiple linear regression models; scores adjusted for age of the child at time of test, order of the child (between siblings), nursery at 14 months, breastfeeding, maternal social class, IQ, and smoking, and cord blood mercury level					
	Boys		125		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	10.54	0.19–20.89	p<0.05	1 mg increase in maternal urinary fluoride across whole pregnancy associated with a 10.54-point (0.19 to 20.89) higher general cognitive

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
									score in boys when adjusted for blood cord mercury levels; no association for girls. 1mg increase in maternal urinary fluoride in third trimester associated with a 8.15-point (0.69 to 15.61) higher general cognitive score in boys when adjusted for blood cord mercury levels; no association for girls.
	Girls		123		Maternal urinary fluoride (whole pregnancy); creatinine-adjusted	-0.83	-8.18–6.52	p>0.05	
	Total sample		248		Maternal urinary fluoride (first trimester); creatinine-adjusted	1	-4.61–6.61	p>0.05	
	Boys		125		Maternal urinary fluoride (third trimester); creatinine-adjusted	8.15	0.69–15.61	p<0.05	
	Girls		123		Maternal urinary fluoride (third trimester);	-0.46	-6.04–5.12	p>0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
					creatinine-adjusted				
Hyperactivity/inattention scores									
Riddell *et al.* (2019) (CHMS) ●		Scores, hyperactivity/inattention scale on the Strengths and Difficulties Questionnaire (Goodman 2001)							
	Urinary fluoride sample		1877	Linear regression models; scores adjusted for child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	Specific gravity-adjusted urinary fluoride	0.31	-0.04–0.66	p>0.05	Urinary fluoride did not significantly predict hyperactivity/inattention scale scores
	Fluoride in tap water sample		710	Linear regression models; scores adjusted for child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education,	Fluoride concentration in tap water	0.31	0.04–0.58	p<0.05	1mg/L increase in tap water fluoride was associated with a 1.52 (0.23–2.80) increase in HI scores for youth in 75th percentile; not significant for youth in 25th percentile

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
				total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level					
	Fluoride in tap water sample - 75th age percentile					1.52	0.23–2.80	p<0.05	
	Fluoride in tap water sample - 25th age percentile					-0.33	-1.51–0.84	p>0.05	
	CWF status sample		1722	Linear regression models; scores adjusted for child's sex, age at interview, ethnicity (white or other), body mass index, highest level of parental education, total household income, exposure to cigarette smoke inside the home (yes/no), concurrent blood lead level	CWF status	0.11	0.02–0.20	p<0.05	Significant interaction between age and hyperactivity/inattention scale scores, such that for older youth (75th percentile), living in fluoridated region was associated with a 0.7-point higher score (0.34–1.06); no association between CWF status and scores for youth in 25th percentile
	CWF status sample -					0.7	0.34–1.06	p<0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Significance	Summary of findings
	75th age percentile								
	CWF status sample - 25th age percentile					0.04	-0.38–0.46	p>0.05	

Table 81 Findings for ecological studies examining bone cancer (osteosarcoma, Ewing sarcoma, bone cancer)

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
Osteosarcoma									
Hrudey *et al.* (1990)		Incidence of osteosarcoma recorded by the Alberta Cancer Board from 1970 to 1988		Crude rate per 100,000 per year calculated	Crude rate per 100,000 per year		-		
	Edmonton (fluoridated)		26		0.27	-	-	-	Similar incidence of osteosarcoma in fluoridated and non-fluoridated communities (no statistical comparison performed)
	Calgary (non-fluoridated)		29		0.29	-			
Mahoney *et al.* (1991)		Incidence of osteosarcoma recorded by the New York		Crude rate per 100,000 per year calculated.			Standard error		No difference in incidence between areas with and without fluoridated

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		State Cancer Registry from 1975 to 1987		standard error estimated by dividing rate by square root of the number of cases on which it was based					water for any comparisons
	Men non-fluoridated <30 years		34		0.44 †	-	0.075 †		
	Men fluoridated Standard Metropolitan Statistical Area <30 years		75		0.43 †	-	0.043 †		
	Men fluoridated non-Standard Metropolitan Statistical Area <30 years		26		0.49 †	-	0.096 †		
	Women non-fluoridated <30 years		24		0.32 †	-	0.064 †		
	Women fluoridated Standard Metropolitan Statistical Area <30 years		44		0.25 †	-	0.038 †		
	Women fluoridated non-Standard		10		0.2 †	-	0.062 †		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Metropolitan Statistical Area <30 years								
	Men non-fluoridated >30 years		26		0.33 †	-	0.064 †		
	Men fluoridated Standard Metropolitan Statistical Area >30 years		50		0.29 †	-	0.041 †		
	Men fluoridated non-Standard Metropolitan Statistical Area >30 years		14		0.29 †	-	0.077 †		
	Women non-fluoridated >30 years		22		0.24 †	-	0.051 †		
	Women fluoridated Standard Metropolitan Statistical Area >30 years		45		0.22 †	-	0.033 †		
	Women fluoridated non-Standard Metropolitan Statistical Area >30 years		8		0.15 †	-	0.051 †		
Cohn (1992)		Incidence of osteosarcoma in people aged		Relative rate ratios, CIs calculated	Relative rate ratio		95% CI		Rate ratios elevated in fluoridated areas for males between

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		under 20 years compiled from the New Jersey Cancer Registry between 1979 and 1987		using tables from Haenszel *et al.* (1962)					10-19 years. No difference in rate ratios for females or men in older age groups
	Male - fluoridated - <20 years		12		3.4		1.8–6.0		
	Male - non-fluoridated <20 years		8						
	White male - fluoridated - 10-19 years		-		4.8		2.3–8.8		
	White male - non-fluoridated - 10-19 years		-						
	Male - 0-9 - fluoridated		2						
	Male - 0-9 - non-fluoridated		1						
	Male - 10-19 - fluoridated		10		3.4		1.7–6.4		
	Male - 10-19 - non-fluoridated		7						
	Male - 20-49 - fluoridated		5		2.6				
	Male - 20-49 - non-fluoridated		5						
	Male - 50-69 - fluoridated		0		-				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Male - 50-69 - non-fluoridated		7						
	Male - 70+ - fluoridated		1		0.6 †				
	Male - 70+ - non-fluoridated		4						
	Female - 0-9 - fluoridated		0		-				
	Female - 0-9 - non-fluoridated		2						
	Female - 10-19 - fluoridated		3		1.4 †				
	Female - 10-19 - non-fluoridated		5						
	Female - 20-49 - fluoridated		2		0.9 †				
	Female - 20-49 - non-fluoridated		5						
	Female - 50-69 - fluoridated		1		1.2 †				
	Female - 50-69 - non-fluoridated		2						
	Female - 70+ - fluoridated		5		2.6 †				
	Female - 70+ - non-fluoridated		4						
Comber *et al.* (2011)		Osteosarcoma incidence in Northern Ireland and Republic of Ireland 1994-2006		Rate ratio for osteosarcoma incidence, Republic of Ireland fluoridated compared with all-	Standardised incident rate ratio		95% CI		No evidence of a significant association between water fluoridation and osteosarcoma incidence for any age or sex groups

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				Ireland non-fluoridated					
	Females 0-24 years		-		1.05	Age	0.83–1.33	-	
	Females all ages		-		1.34	Age	0.83 –2.17	-	
	Males 0-24 years		-		1	Age	0.85–1.17	-	
	Males all ages		-		1.09	Age	0.75–1.59	-	
	Total 0-24 years		-		1.01	Age	0.88–1.15	-	
	Total all ages		-		1.17	Age	0.87–1.58	-	
National Fluoridation Information Service (2013)		Osteosarcoma incidence, diagnosed 2000-2008		No inferential statistical analysis specified, only rate per million calculated	Risk ratio				Osteosarcoma is extremely rare in New Zealand, with an average of 14.1 cases per year. No difference in rates of osteosarcoma cases between areas with and without CWF for both sexes (no statistical comparisons performed; descriptive statistics only).
	Male 0 - 9 CWF		2		0.302 †	-	-		
	Male 0 - 9 Non-CWF		5			-	-		
	Male 10-19 CWF		17		0.835 †	-	-		
	Male 10-19 Non-CWF		16			-	-		
	Male 20 - 39 CWF		8		0.722 †	-	-		
	Male 20 - 39 Non-CWF		7			-	-		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Male 40 - 64 CWF		7		1.278 †	-	-		
	Male 40 - 64 Non-CWF		5			-	-		
	Male 65+ CWF		4		0.606 †	-	-		
	Male 65+ Non-CWF		7			-	-		
	Male All ages CWF		38		-	-	-		
	Male All ages Non-CWF		40			-	-		
	Female 0 - 9 CWF		3		2.333 †	-	-		
	Female 0 - 9 Non-CWF		1			-	-		
	Female 10-19 CWF		16		1.097 †	-	-		
	Female 10-19 Non-CWF		11			-	-		
	Female 20 - 39 CWF		3		0.900 †	-	-		
	Female 20 - 39 Non-CWF		2			-	-		
	Female 40 - 64 CWF		9		-	-	-		
	Female 40 - 64 Non-CWF		0			-	-		
	Female 65+ CWF		0		0.000 †	-	-		
	Female 65+ Non-CWF		4			-	-		
	Female All ages CWF		31		-	-	-		
	Female All ages Non-CWF		18			-	-		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Both sexes 0 - 9 CWF		5		0.630 †	-	-		
	Both sexes 0 - 9 Non-CWF		6			-	-		
	Both sexes 10-19 CWF		33		0.937 †	-	-		
	Both sexes 10-19 Non-CWF		27			-	-		
	Both sexes 20 - 39 CWF		11		0.739 †	-	-		
	Both sexes 20 - 39 Non-CWF		9			-	-		
	Both sexes 40 - 64 CWF		16		2.778 †	-	-		
	Both sexes 40 - 64 Non-CWF		5			-	-		
	Both sexes 65+ CWF		4		0.354 †	-	-		
	Both sexes 65+ Non-CWF		11			-	-		
	Both sexes All ages CWF		69			-	-		
	Both sexes All ages Non-CWF		58			-	-		
Blakey *et al.* (2014)		Osteosarcoma incidence, diagnosed from 1980 to 2005 in those aged 0–49 years		Negative binomial regression, independent variables were census-derived SAU attributes allocated to	Risk ratio for 1ppm increase in fluoride level (adjusted)		95% CI		No association between artificial fluoridation (as a binary variable with adjustment for deprivation) and osteosarcoma

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
				2001 census geography					
	Cases 1970-1979				1.046	Compared with the non-fluoride model that contained age-group, gender, gender*age-group, non-home ownership, non-home ownership *age-group plus adjustment for deprivation—cohort is restricted to include cases born between 1970 and 1979	0.870–1.257	p>0.05	
	Cases 1980 or later				1.218	Compared with the non-fluoride model that contained age-group, gender, gender*age-group, unemployment, unemployment*age-group plus adjustment for deprivation—cohort is restricted to include cases born 1980 or later	1.021–1.452	p>0.05	
Young *et al.* (2015)		All cases in England diagnosed between 1995 and 2010		Incidence rate ratio	Incidence rate ratio		95% CI		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
		recorded in cancer registries							
	All under-25s - Fluoridated areas		148		8.2	Age, gender, deprivation, ethnicity	-9.3–29	p=0.38	No difference in rate of osteosarcoma in under-25s between fluoridated and non-fluoridated areas following adjustment for age, gender, deprivation, and ethnicity
	All under-25s - Non-fluoridated areas		949						
	Male under-25s - Fluoridated areas		82		17	Age, gender, deprivation, ethnicity	-7.1–46	p=0.19	No difference in rate of osteosarcoma in male under-25s between fluoridated and non-fluoridated areas following adjustment for age, gender, deprivation, and ethnicity
	Male under-25s - Non-fluoridated areas		540						
	Female under-25s - Fluoridated areas		56		-2.5	Age, gender, deprivation, ethnicity	-27–30	p=0.86	No difference in rate of osteosarcoma in female under-25s between fluoridated and non-fluoridated areas following adjustment for age,

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
									gender, deprivation, and ethnicity
	Female under-25s - Non-fluoridated areas		409						
	All over 50s - Fluoridated areas		73		-15	Age, gender, deprivation, ethnicity	-34–9.6	p=0.21	No difference in rate of osteosarcoma in over-50s between fluoridated and non-fluoridated areas following adjustment for age, gender, deprivation, and ethnicity
	All over 50s - Non-fluoridated areas		587						
Ewing sarcoma									
Blakey *et al.* (2014)		Ewing sarcoma incidence, diagnosed from 1980 to 2005, in those aged 0–49 years		Negative binomial regression, independent variables were census-derived SAU attributes allocated to 2001 census geography	Risk ratio for 1ppm increase in fluoride level (adjusted)		95% CI		
	Cases 1970-1979		-		0.987	Compared with the non-fluoride model that contained age-group, gender, age-	0.796–1.223	p>0.05	No association between artificial fluoridation (as a binary variable with

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
						group *gender, population density, unemployment—cohort is restricted to include cases born between 1970 and 1979.			adjustment for deprivation) and Ewing sarcoma
	Cases 1980 or later		-		0.999	Compared with the non-fluoride model that contained age-group, gender, age-group *gender, Scotland, unemployment—cohort is restricted to include cases born 1980 or later.	0.796–1.255	p=0.996	No association between artificial fluoridation (as a binary variable with adjustment for deprivation) and Ewing sarcoma
Bone cancer									
Mahoney *et al.* (1991)		Incidence of bone cancer recorded by the New York State Cancer Registry from 1975 to 1987		Crude rate per 100,000 per year calculated. standard error estimated by dividing rate by square root of the number of cases on which it was based			Standard error		No difference in incidence between areas with and without fluoridated water for any comparisons
	Men non-fluoridated <30 years		75		0.97 †	-	0.112 †		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Men fluoridated Standard Metropolitan Statistical Area <30 years		158		0.09 †	-	0.071 †		
	Men fluoridated non-Standard Metropolitan Statistical Area <30 years		60		1.13 †	-	0.146 †		
	Women non-fluoridated <30 years		98		0.67 †	-	0.094 †		
	Women fluoridated Standard Metropolitan Statistical Area <30 years		181		0.64 †	-	0.061 †		
	Women fluoridated non-Standard Metropolitan Statistical Area <30 years		44		0.43 †	-	0.092 †		
	Men non-fluoridated >30 years		51		1.24 †	-	0.125 †		
	Men fluoridated Standard Metropolitan Statistical Area >30 years		110		1.04 †	-	0.078 †		

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Men fluoridated non-Standard Metropolitan Statistical Area >30 years		22		0.9 †	-	0.136 †		
	Women non-fluoridated >30 years		64		0.7 †	-	0.087 †		
	Women fluoridated Standard Metropolitan Statistical Area >30 years		162		0.8 †	-	0.062 †		
	Women fluoridated non-Standard Metropolitan Statistical Area >30 years		38		0.67 †	-	0.112 †		
Lee *et al.* (2020)		Incidence of bone cancer, data gathered from National Health Insurance Scheme		Bayesian spatio-temporal regression analysis to calculate posterior relative risk	Relative risk		95% Credible Interval		
	Total sample		-		1.2	-	0.89–1.61	p>0.05	Relative risks increased over time but did not increase in CWF area compared to non-CWF areas

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	Variability statistic	Significance	Summary of findings
	Male		-		1.26	-	0.84–1.88	p>0.05	
	Female		-		1.03	-	0.87–1.22	p>0.05	

Note: † denotes figures that were not extracted directly from papers but calculated by the review team based on provided information.

Table 82 Findings for case-control studies examining osteosarcoma

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Osteosarcoma									
McGuire *et al.* (1991)		Cases of osteosarcoma (before the age of 40 years), diagnosed from 1980 to 1990							No associations were significant; no evidence that exposure to fluoride is a risk factor for osteosarcoma
	Pairs in which only case lived more than 1/3 life at >0.7ppm Pairs in which only control lived more than 1/3 life at >0.7ppm				0.14		0.02–1.22	p>0.05	
	Pairs in which only case had high average exposure Pairs in which only control had high average exposure				0.33		0.04–2.5	p>0.05	
	Pairs in which only case lived more than 1/3 of first 15 years of life at >0.7ppm				0.33		0.04–2.5	p>0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Pairs in which only control lived more than 1/3 of first 15 years of life at >0.7ppm				0.14		0.02–1.22	p>0.05	
Bassin *et al.* (2006) (Harvard)		Cases of osteosarcoma (before the age of 20 years)							
	Male Less than 30% of target fluoride exposure at age 7 years (reference)		-	Conditional logistic regression, using less than 30% of target fluoride exposure group as reference	Reference	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements			For males diagnosed with osteosarcoma before age 20, fluoride level in drinking water before age 7 was associated with higher risk of osteosarcoma
	Male 30-99% of target fluoride exposure at age 7 years		-		3.36	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements	0.99–11.42	-	
	Male At least 100% of target fluoride exposure at age 7 years		-		5.46	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements	1.50–19.90	-	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Female Less than 30% of target fluoride exposure at age 7 years (reference)		-	Conditional logistic regression, using less than 30% of target fluoride exposure group as reference	Reference	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements			For females diagnosed with osteosarcoma before age 20, no association between fluoride level in drinking water before age 7 and risk of osteosarcoma
	Female 30-99% of target fluoride exposure at age 7 years		-		1.39	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements	0.41–4.76	-	
	Female At least 100% of target fluoride exposure at age 7 years		-		1.75	Age, zip code median income, county population, use of well water by age 7, use of bottled water by age 7, any use of fluoride supplements	0.48–6.35	-	
Kim *et al.* (2020) (Harvard)		Incidence of osteosarcoma in participating departments from 1989 to 1993 and 1994 to 2000		Multivariate logistic regression	Adjusted odds ratio				
	Ever lived in fluoridated area - No bottled water		437		0.51	race, ethnicity, income, ever lived in urban residence, distance from hospital	0.31–0.84	p=0.008	Ever having lived in a fluoridated community, including those who did not drink bottled water,

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
									showed significant protective effect with osteosarcoma; not demonstrated for those who drank bottled water
	Never lived in fluoridated area - No bottled water		101		Ref				
	Ever lived in fluoridated area - Bottled water		69		1.86	race, ethnicity, income, ever lived in urban residence, distance from hospital	0.54–6.41	p=0.326	
	Never lived in fluoridated area - Bottled water		38		Ref				
	0% life lived in fluoridated area		139		Ref				Significant protective effect for those who lived <50% of their lives in fluoridated community compared to those who never did; no effect for those who lived >50% lives compared with those who did not
	<50% life lived in fluoridated area		108		0.41	race, ethnicity, income, ever lived in urban residence, distance from hospital, and ever drank bottled water	0.22–0.76	p=0.005	
	50-100% life lived in fluoridated area		132		0.69	race, ethnicity, income, ever lived in urban residence, distance from	0.40–1.21	p=0.198	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	100% life lived in fluoridated area		266		0.67	hospital, and ever drank bottled water race, ethnicity, income, ever lived in urban residence, distance from hospital, and ever drank bottled water	0.38–1.18	p=0.163	

Table 83 Findings for studies examining other cancers

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Secondary bone cancer									
Crnosija *et al.* (2019)		County-level percentage of secondary bone cancer over cancer diagnosis		Ordinary least squares regression, one predictor variable: categorised percentage of population receiving fluoridated water, using <25% population in county with fluoridation as reference	B coefficient				
	<25% population in county with		-		Reference	Age	-	-	No relationship between county-level % access to fluoridated

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	fluoridation (reference)								water and prevalence of secondary bone cancer diagnosis among cancer patients
	25-75% population in county with fluoridation		-		0.02		-	p=0.96	
	>75% population in county with fluoridation		-		0.02		-	p=0.97	
Bladder cancer									
Young *et al.* (2015)		All primary invasive bladder cancer cases in England diagnosed between 2000 and 2010 and recorded in cancer registries		Incidence rate ratio	Incidence rate ratio (adjusted)				
	Fluoridated areas		11,327		-8	Age, gender, deprivation, ethnicity	-9.9–-6.0	p<0.001	Lower rates of bladder cancer in fluoridated areas
	Non-fluoridated areas		84,780						
All cancers									
Young *et al.* (2015)		All cancer cases in England (excluding non-melanoma		Incidence rate ratio	Incidence rate ratio (adjusted)				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
		skin cancer) diagnosed between 2007 and 2010 and recorded in cancer registries							
	Fluoridated areas		131,288		-0.4	Age, gender, deprivation, ethnicity	-1.2–0.4	p=0.029	All cancer incidence lower in fluoridated areas following adjustment for age, gender, deprivation; however, this was not maintained when also adjusted for ethnicity
	Non-fluoridated areas		921,583						

Table 84 Findings for studies examining endocrine conditions

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Hypothyroidism prevalence									
Peckham *et al.* (2015)		Practice-level hypothyroidism prevalence	946 practices across both areas	Binary logistic regression to predict likelihood of practice being categorised as recording high level of hypothyroidism,	Odds ratio for upper tertile hypothyroidism prevalence				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
				using non-fluoridated area as reference					
	Fluoridated area (West Midlands)		-		1.935	Proportion of women registered with general practice, proportion of patients over 40 years old registered with general practice, Index of Multiple Deprivation categories	1.388 – 2.699	P<0.001	Practices in fluoridated area (West Midlands) have nearly twice the odds of recording a high level of hypothyroidism compared to practices in non-fluoridated area (Greater Manchester)
	Non-fluoridated area (Greater Manchester) (reference)		-						
Diagnosed thyroid condition									
Barberio *et al.* (2017b) (CHMS) ●		Self-reported diagnosis of a thyroid condition, yes/no			Odds ratio				
	Cycle 2 respondents aged 12 and over, fluoride		-	Logistic regression, outcome	0.98	age, sex, household income adequacy,	0.95–1.02		No association between fluoride exposure

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	urine subsample			regressed on urinary fluoride		highest attained education in the household			measured by urinary fluoride and self-reported diagnosis of a thyroid condition
	Cycle 3 respondents aged 12 and over, fluoride urine subsample		-	Logistic regression, outcome regressed on urinary fluoride	1.00	age, sex, household income adequacy, highest attained education in the household	0.99–1.01		No association between fluoride exposure measured by urinary fluoride and self-reported diagnosis of a thyroid condition
	Cycle 3 respondents aged 12 and over, fluoride tap water subsample		-	Logistic regression, outcome regressed on fluoride concentration of tap water	0.98	age, sex, household income adequacy, highest attained education in the household	0.28–3.45		No association between fluoride exposure measured by tap water and self-reported diagnosis of a thyroid condition
Low TSH levels									
Barberio *et al.* (2017b) (CHMS) ●		Blood samples collected by a phlebotomist using standardised venepuncture method, quantification of TSH in serum determined using third-generation assay analyser with			Odds ratio				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
		a chemiluminescent detection system							
	Cycle 3 respondents aged 12 and over, fluoride urine subsample		-	Logistic regression, outcome regressed on urinary fluoride, reference group normal TSH	1.01	age, sex, household income adequacy, highest attained education in the household	0.99–1.04	p>0.05	No association between fluoride exposure measured by urinary fluoride and low TSH levels
	Cycle 3 respondents aged 12 and over, fluoride tap water subsample		-	Logistic regression, outcome regressed on fluoride concentration of tap water, reference group normal TSH	1.38	age, sex, household income adequacy, highest attained education in the household	0.08–24.49	p>0.05	No association between fluoride exposure measured by tap water and low TSH levels
High TSH levels									
Barberio *et al.* (2017b) (CHMS)		Blood samples collected by a phlebotomist using standardised venipuncture method, quantification of TSH in serum determined using third-generation assay analyser with a chemiluminescent detection system			Odds ratio				
	Cycle 3 respondents aged 12 and over, fluoride		-	Logistic regression, outcome regressed on	0.99	age, sex, household income adequacy,	0.97–1.02	p>0.05	No association between fluoride exposure

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	urine subsample			urinary fluoride, reference group normal TSH		highest attained education in the household			measured by urinary fluoride and high TSH levels
	Cycle 3 respondents aged 12 and over, fluoride tap water subsample		-	Logistic regression, outcome regressed on fluoride concentration of tap water, reference group normal TSH	1.2	age, sex, household income adequacy, highest attained education in the household	0.14–10.08	p>0.05	No association between fluoride exposure measured by tap water and high TSH levels
Less than recommended sleep duration									
Cunningham *et al.* (2021) (CHMS) ●		Self-reported habitual sleep duration, reported to the closest half-hour, categorised as lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups			Adjusted relative risk ratio association for change in outcome per 0.5mg/L change in tap water fluoridation/specific gravity-adjusted urinary fluoride concentration, using recommended sleep duration as reference				
	Water fluoride sample		1,016	Multinomial logistic regression of sleep duration on water	1.34	BMI, ethnicity, total household income,	1.03 – 1.73	p<0.05	For every 0.5 mg/L increase in water fluoride concentration,

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
				fluoride levels adjusted for covariates		chronic health condition yes/no, sex, age			there was a 34% increased relative risk of reporting sleeping less than the recommended duration
	Urinary fluoride sample		1,303	Multinomial logistic regression of sleep duration on specific gravity-adjusted urinary fluoride levels adjusted for covariates	1.02	BMI, ethnicity, total household income, chronic health condition yes/no, sex, age	0.93 – 1.13	p=0.64	No association between change in urinary fluoride and relative risk of sleeping less than recommended duration
More than recommended sleep duration									
Cunningham *et al.* (2021) (CHMS)		Self-reported habitual sleep duration, reported to the closest half-hour, categorised into lower than recommended/recommended/higher than recommended based on National Sleep Foundation sleep range recommendations for relevant age groups			Adjusted relative risk ratio association for change in outcome per 0.5mg/L change in tap water fluoridation/specific gravity-adjusted urinary fluoride concentration, using recommended sleep duration as reference				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Water fluoride sample		1,016	Multinomial logistic regression of sleep duration on water fluoride levels adjusted for covariates	0.95	BMI, ethnicity, total household income, chronic health condition yes/no, sex, age	0.53 – 1.71	p>0.05	No association between change in fluoride concentration and relative risk of sleeping more than recommended duration
	Urinary fluoride sample		1,303	Multinomial logistic regression of sleep duration on specific gravity-adjusted urinary fluoride levels adjusted for covariates	0.91	BMI, ethnicity, total household income, chronic health condition yes/no, sex, age	0.74 – 1.13	p=0.40	No association between change in urinary fluoride and relative risk of sleeping more than recommended duration
Trouble sleeping									
Cunningham *et al.* (2021) (CHMS) ●		Self-reported frequency of sleep problems, single question with five-point response scale			Adjusted odds ratio association for change in outcome per 0.5mg/L change in tap water fluoridation/specific gravity-adjusted urinary fluoride concentration				
	Water fluoride sample		1,016	Multinomial logistic regression of trouble sleeping	0.95	BMI, ethnicity, total household	0.77, 1.18	p=0.67	No association between change in fluoride

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
				on water fluoride levels adjusted for covariates		income, chronic health condition yes/no, sex, age			concentration and relative risk of trouble sleeping
	Urinary fluoride sample		1,303	Multinomial logistic regression of trouble sleeping on specific gravity-adjusted urinary fluoride levels adjusted for covariates	0.96	BMI, ethnicity, total household income, chronic health condition yes/no, sex, age	0.89, 1.04	p=0.37	No association between change in urinary fluoride and relative risk of trouble sleeping
Daytime sleepiness									
Cunningham *et al.* (2021) (CHMS) ●		Self-reported frequency of daytime sleepiness, single question with five-point response scale			Adjusted odds ratio association for change in outcome per 0.5mg/L change in tap water fluoridation/specific gravity-adjusted urinary fluoride concentration				
	Water fluoride sample		1,016	Multinomial logistic regression of daytime sleepiness on water fluoride levels adjusted for covariates	1.16	BMI, ethnicity, total household income, chronic health condition	0.94, 1.44	p=0.17	No association between change in fluoride concentration and relative risk of daytime sleepiness

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
						yes/no, sex, age			
	Urinary fluoride sample		1,303	Multinomial logistic regression of daytime sleepiness on specific gravity-adjusted urinary fluoride levels adjusted for covariates	1	BMI, ethnicity, total household income, chronic health condition yes/no, sex, age	0.92, 1.08	p=0.95	No association between change in urinary fluoride and relative risk of daytime sleepiness

Table 85 Findings for studies examining endocrine conditions using regression analysis

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Standard error	Significance	Summary of findings
TSH levels										
Malin *et al.* (2018) (CHMS)	Total sample	Serum TSH levels measured using a third-generation assay analyser equipped with a chemiluminescent detection system	-	Linear regression, modelling serum TSH levels as a function of urinary fluoride and iodine levels controlling for covariates (age, BMI, serum						No evidence for an association between urinary fluoride and TSH in the absence of iodine status; however, moderate-to-severe iodine deficiency revealed an association of a 0.35mIU/L increase in TSH for every 1mg/L increase in urinary fluoride (specific gravity

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Variable	B coefficient	95% CI	Standard error	Significance	Summary of findings
				calcium and sex)						adjusted) (interaction)
					Specific gravity-adjusted urinary fluoride	-0.02	-0.19–0.15	0.09	p=0.43	
					Urinary iodine	-0.55	-0.80–-0.31	0.12	p=0.00	
					Specific gravity-adjusted urinary fluoride * urinary iodine	0.36	-0.03–0.75	0.2	p=0.03	

Table 86 Findings for studies examining renal conditions

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Incidence rate ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Renal calculi									
Young *et al.* (2015)		Incidence of kidney stones inpatient consultant episodes per lower super output area in England recorded in hospital episode statistics; admission dates between	-	Incidence rate ratio					

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Incidence rate ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
		April 2007 and March 2013							
	Fluoridated areas		18,579		-7.9	Age, gender, deprivation, ethnicity	-9.6–-6.2	p<0.001	Strong evidence that rate of kidney stones was lower in fluoridated areas compared to non-fluoridated areas, controlling for confounders
	Non-fluoridated areas		141,963						

Table 87 Findings for studies examining birth and birthing abnormalities

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Down syndrome									
Lowry *et al.* (2003)		Cases of Down syndrome, all cases of a congenital abnormality with a final postnatal diagnosis of trisomy 21 using International Classification of Diseases, 9th Revision (ICD-9) codes	-	Generalised linear modelling with a Poisson error structure and log link function	Odds ratio				
	Fluoridated areas				1.05	Material deprivation, district of residence, maternal age	0.79–1.41		No significant association between water fluoridation and incidence of Down syndrome

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Non-fluoridated areas								
Young *et al.* (2015)		Cases of Down syndrome according to lower-tier local authority, recorded in the National Down Syndrome Cytogenetic Register, including live births, stillbirths, late miscarriages, and terminations of pregnancy with foetal anomaly, from 2009 to 2012 inclusive	-	Poisson regression model, adjusted for total number of births	Odds ratio				
	Fluoridated areas		658		1.7	Maternal age	-6.2–10	p=0.68	No evidence of a difference in rate of Down syndrome in fluoridated and non-fluoridated areas when adjusted for maternal age
	Non-fluoridated areas		5,961						
All trisomies									
Lowry *et al.* (2003)		All cases of a congenital abnormality with a final postnatal diagnosis of a trisomy (trisomies 21, 13, and 18 only, ICD-9 codes 758.0, 758.1, and 758.2)		Generalised linear modelling with a Poisson error structure and log link function	Odds ratio				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Fluoridated areas		-		1.11	Material deprivation, district of residence, maternal age	0.86–1.43	p>0.05	No significant association between water fluoridation and incidence of trisomies
	Non-fluoridated areas		-						
Neural tube defects									
Lowry *et al.* (2003)		All cases of a congenital abnormality with a final postnatal diagnosis of a neural tube defect (as defined by the EUROCAT (European Registration of Congenital Anomalies and Twins) system of classification, ICD-9 codes 740.0, 740.1, 740.2, 741.0, 741.9, and 742.0)		Generalised linear modelling with a Poisson error structure and log link function	Odds ratio				
	Fluoridated areas		-		0.82	Material deprivation, district of residence, maternal age	0.62–1.09	p>0.05	No significant association between water fluoridation and incidence of neural tube defect
	Non-fluoridated areas		-						
Clefts									
Lowry *et al.* (2003)		All cases of a congenital abnormality with a final postnatal diagnosis of a facial cleft (cleft palate,		Generalised linear modelling with a Poisson error	Odds ratio				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
		cleft lip with or without cleft palate, Pierre Robin syndrome, ICD-9 codes 749.0, 749.1,749.2, 756.03)		structure and log link function					
	Fluoridated areas		-		0.63	Material deprivation, district of residence, maternal age	0.46–0.86	p>0.05	No significant association between water fluoridation and incidence of clefts
	Non-fluoridated areas		-						
Stillbirths									
Lowry *et al.* (2003)		All stillbirths between 1 January 1989 and 31 December 1998 identified from the Northern Perinatal Mortality Survey		Generalised linear modelling with a Poisson error structure and log link function	Odds ratio				
	Fluoridated areas		-		1.06	Material deprivation, district of residence, maternal age	0.91, 1.24	p>0.05	No significant association between water fluoridation and incidence of stillbirth
	Non-fluoridated areas		-						
Preterm births									
Zhang *et al.* (2019)		Incidence of pre-term births (<37 weeks)		Adjusted risk ratios, using neither dental	Risk ratio				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Summary statistic (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
				cleaning nor CWF as reference					
	Neither dental cleaning nor CWF (reference)		1,283		1	Maternal race/ethnicity, age, education, federal poverty line, insurance status, pre-pregnancy diabetes, prior preterm birth, nativity, and pre-pregnancy BMI			Women with dental cleaning alone and dental cleaning plus CWF have significantly lower incidence of preterm births compared with those who had neither, after controlling for confounding. No impact of CWF alone on preterm births.
	Dental cleaning only		1,614		0.74		0.55–0.98	p<0.05	
	CWF only		2,908		0.81		0.63–1.05	p>0.05	
	Dental cleaning and CWF		3,429		0.74		0.57–0.95	p<0.05	

Table 88 Findings for studies examining infant abnormalities

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Odds ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
Sudden infant death syndrome									
Dick *et al.* (1999)		Cases of SIDS drawn from the New Zealand Cot		Odds ratios, calculated from relative risks calculated via					

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Odds ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
		Death Study, 1987-1990		logistic regression					
	New Zealand - non-fluoridated (reference)		152 SIDS, 606 control	Odds ratio using non-fluoridated infants as reference group	1	Infant age, region, time, season, occupational status, marital status, age mother left school, age of mother at infant's birth, age of mother at first pregnancy, number of previous pregnancies, smoking status, alcohol intake over last trimester, caffeine consumption in third trimester, attendance at antenatal clinics and classes, maternal weight, infant sex, birthweight, gestation, ethnicity, twin, admitted to neonatal unit, method of feeding at discharge, sleep position, bed sharing			Prenatal exposure to fluoridated water was not associated with higher risk for SIDS in either the full sample or the North Island sample (which excluded the South Island, where fluoride supplementation was more prevalent at the time of the study).
	New Zealand - fluoridated		227 SIDS, 944 control		1.19		0.82–1.74	p>0.05	
	North Island - non-fluoridated (reference)		87 SIDS, 370 control	Odds ratio using non-fluoridated infants as reference group	1				

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Odds ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	North Island - fluoridated		184 SIDS, 837 control		1.06		0.69–1.64	p>0.05	
	Non-fluoridated - breast feeding (reference)		67 SIDS, 361 control	Odds ratio using non-fluoridated breastfed infants as reference group	1	Infant age, region, time, season, occupational status, marital status, age mother left school, age of mother at infant's birth, age of mother at first pregnancy, number of previous pregnancies, smoking status, alcohol intake over last trimester, caffeine consumption in third trimester, attendance at antenatal clinics and classes, maternal weight, infant sex, birthweight, gestation, ethnicity, twin, admitted to neonatal unit, sleep position, bed sharing			No higher risk of SIDS with breast feeding in fluoridated areas; no evidence for interaction between fluoridated water supplies and infant feeding.
	Fluoridated - breast feeding		94 SIDS, 554 control		1.09		0.66–1.79	p>0.05	
	Non-fluoridated–formula		67 SIDS, 178 control		1.38		0.83–2.28	p>0.05	

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Odds ratio (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
	Fluoridated - formula		101 SIDS, 272 control		1.71		1.02–2.90	p<0.05	
	Non-fluoridated – formula (reference)		67 SIDS, 178 control	Odds ratio using non-fluoridated infants as reference group	1	Infant age, region, time, season, occupational status, marital status, age mother left school, age of mother at infant's birth, age of mother at first pregnancy, number of previous pregnancies, smoking status, alcohol intake over last trimester, caffeine consumption in third trimester, attendance at antenatal clinics and classes, maternal weight, infant sex, birthweight, gestation, ethnicity, twin, admitted to neonatal unit, sleep position, bed sharing			Fluoridated formula feeding showed no higher risk of SIDS compared with non-fluoridated formula feeding
	Fluoridated - formula		101 SIDS, 272 control		1.25		0.73–2.13	p>0.05	

Table 89 Findings for studies examining all-cause mortality

Paper	Participant group	Units, method of measurement	Number of cases	Statistical method of analysis	Incidence rate ratio (%) (adjusted)	Adjusted for…	95% CI	Significance	Summary of findings
All-cause mortality									
Young *et al.* (2015)		Count of deaths, obtained at lower super output area level from Office of National Statistics, in England for January 2009-January 2012		Incidence rate ratio, adjusted					
	Fluoridated areas		233,922		-1.3	Age, gender, deprivation, ethnicity	-2.5–-0.1	p=0.04	Some evidence that rate of deaths from all recorded causes was lower in fluoridated areas than non-fluoridated areas, but effect size was small.
	Non-fluoridated areas		1,602,206						

Appendix O Summary of excluded studies of natural fluoridation

Purpose

Our reasons for excluding studies of areas with naturally fluoridated water within WHO guideline limits are presented in Section 2.2. While the decision to exclude these studies has led to the loss of some data on certain outcomes (e.g. diabetes), it has allowed for a much more specific and more appropriate analysis with a tightly defined exposure that has direct relevance to policy decisions being made in Ireland. Furthermore, based on the screening process from this review (see Appendix C for a list of studies excluded on full text screening) and the findings of the HRB's 2015 review [25], which did include studies of areas with natural fluoridation within recommended levels, we believed that the studies excluded on this basis would not substantially alter the conclusions of the review on artificial water fluoridation.

However, we acknowledge that the findings from studies of natural fluoridation within the WHO guideline limits may offer additional useful data for understanding the impact of fluoride. Therefore, this appendix presents a brief overview of additional studies that were excluded from the synthesis due to the presence of natural fluoride within recommended levels in the analysis.

The information presented here is intended as a brief overview only and is not incorporated into the main synthesis. Only high-level extraction was carried out by a single author, and no quality assessment was undertaken on these papers.

Identifying evidence

After the main synthesis was completed, records that had been excluded on the basis of intervention, at either title and abstract or full text stage, were rescreened by one of two screeners. The same inclusion criteria were used as for the main synthesis, with the exception that studies of natural fluoridation or mixed natural and artificial fluoridation were now eligible for inclusion, provided that total fluoride concentration did not exceed 1.5ppm.

A total of 1,747 records were rescreened, yielding 19 papers that met inclusion criteria but had been excluded from the main synthesis due to the inclusion of natural fluoridation in their analysis. An additional study of natural fluoride (Fluegge (2016)), which had been excluded due to the lack of an appropriate comparator, is also discussed here, as it provides useful data on the impact of various fluoridation chemicals in relation to diabetes, which was not examined in any studies included in the main synthesis.

A small number of studies examined the impact of natural fluoridation in areas where the mean fluoride concentration fell between 0.7 and 1.5 ppm, but where the upper range of the concentration exceeded 1.5ppm. These studies were excluded from this supplementary overview and are listed in Table 90.

Table 90 Studies excluded due to upper range of fluoride concentration exceeding 1.5 ppm

Reference
Kurttio P, Gustavsson N, Vartiainen T, Pekkanen J. Exposure to natural fluoride in well water and hip fracture: a cohort analysis in Finland. American journal of epidemiology. 1999 Oct 15;150(8):817-24.
Mustafa DE, Younis UM. The relationship between the fluoride levels in drinking water and the schooling performance of children in rural areas of Khartoum state, Sudan. Fluoride. 2018 Apr 1;51(2):102-13.
Nanayakkara S, Senevirathna ST, Harada KH, Chandrajith R, Nanayakkara N, Koizumi A. The Influence of fluoride on chronic kidney disease of uncertain aetiology (CKDu) in Sri Lanka. Chemosphere. 2020 Oct 1;257:127186.
Shaik N, Shanbhog R, Nandlal B, Tippeswamy HM. Fluoride ingestion and thyroid function in children resident of naturally fluoridated areas-An observational study. Journal of Clinical and Experimental Dentistry. 2019 Oct;11(10):e883.
Wang M, Liu L, Li H, Li Y, Liu H, Hou C, Zeng Q, Li P, Zhao Q, Dong L, Zhou G. Thyroid function, intelligence, and low-moderate fluoride exposure among Chinese school-age children. Environment international. 2020 Jan 1;134:105229.
Hillier S, Cooper C, Kellingray S, Russell G, Hughes H, Coggon D. Fluoride in drinking water and risk of hip fracture in the UK: a case-control study. The Lancet. 2000 Jan 22;355(9200):265-9.
Ding Y, Sun H, Han H, Wang W, Ji X, Liu X, Sun D. The relationships between low levels of urine fluoride on children's intelligence, dental fluorosis in endemic fluorosis areas in Hulunbuir, Inner Mongolia, China. Journal of hazardous materials. 2011 Feb 28;186(2-3):1942-6.

Findings

Brief study characteristics and summarized findings are presented in Table 91.

Bone health

Nine papers, describing nine studies, examined the association between fluoridated water (either natural or mixed natural and artificial fluoridation) and bone characteristics. Five papers examined BMD and five examined incidence of fractures. Eight of the studies found no association or weak associations between fluoridated water and the outcomes of interest. One ecological study found an increased risk of fracture for men only in fluoridated areas.

Neuropsychological outcomes

One prospective cohort study examined the association between prenatal fluoride exposure and measures of cognitive performance at the ages of 4 and 6–12 years. Higher prenatal exposure to fluoride was associated with lower general cognitive scores in children at 4 years and with lower FSIQ scores at 6–12 years; however, sensitivity analysis indicated that associations at 6–12 years may be limited to higher exposures.

Cancer

Five papers, describing five studies, examined the association between fluoridated water (either natural or mixed natural and artificial fluoridation) and incidence of a variety of cancers. One ecological study and one case-control study found no association between fluoridation and osteosarcoma, while one ecological study found that men aged under 20 years in fluoridated areas were at increased risk of bone cancer and osteosarcoma. One ecological study found that states with greater access to fluoridated water had lower incidences of eye cancer. One ecological study found mixed evidence for associations between fluoridation status and a range of cancers; the authors of this study indicated that the consistency of an association between fluoride and cancer was not adequately confirmed by this study.

Endocrine conditions

Three papers, describing two studies, examined the association between fluoridated water (either natural or mixed natural and artificial fluoridation) and endocrine outcomes. Two papers analysed data from the Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) birth cohort study in Mexico City. These papers found that higher exposure to fluoride was associated with later pubertal development for adolescent boys but not girls and increased levels of cardiometabolic risk factors for adolescent girls but not boys.

A 2016 ecological study in the USA by Fluegge found that fluoride artificially added to drinking water to achieve optimal levels (0.7–1.2 ppm) was associated with higher incidence and prevalence of diabetes, when accounting for per capita consumption of tap water and a range of confounders. However, of the three fluoridation chemicals examined as part of the study (sodium fluoride, fluorosilicic acid, or sodium fluorosilicate), fluorosilicic acid was associated with lowest incidence and prevalence of diabetes; this is the chemical used in Ireland for CWF. This is an area requiring further study.

Renal/hepatic outcomes

One cross-sectional study of adolescents in the USA found that higher concentrations of fluoride in plasma were associated with lower estimated glomerular filtration rate, higher serum uric acid concentration, and lower blood urea nitrogen concentration. Higher fluoride concentrations in water were associated with lower blood urea nitrogen concentration. The authors state that the findings are indicative of complex changes in kidney and liver function related to fluoride exposure.

Cardiovascular outcomes

One cohort study in Sweden found no association between fluoride exposure and incidence of myocardial infarction.

Summary

In summary, the findings from the excluded studies of natural fluoride broadly mirror the findings from the main synthesis. Studies of fluoridation and bone health generally returned null or weak findings, as did studies of cancer and cardiovascular outcomes. The findings around neuropsychological and endocrine outcomes are also mirrored in the studies of natural fluoride; studies identified associations between fluoride exposure and endocrine outcomes for adolescents, diabetes, and kidney and liver function, and, at the upper end of guideline levels (1.5 ppm), associated with cognitive performance in children. However, only high-level extraction was carried out on these studies and no quality assessment has been conducted to confirm the reliability of these findings. Our conclusions from the main synthesis stand; further high-quality research is now needed in order to shed light on the impact, if any, of artificial and natural water fluoridation on these aspects of systemic health.

Table 91 Brief study characteristics and summarised findings for excluded studies of natural fluoride

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
Bone health					
Cooper *et al.* (1990)	Ecological or correlational study England and Wales	Comparison of number of hospital admissions for fracturs of proximal femur among men and women aged 45 years and older, 1978–1982, across 39 county districts with varying fluoride concentrations	Natural and artificial fluoride concentrations in 39 county districts ranging 0.0–1.0 ppm	Hip fracture discharge rate	No association between discharge rates and either total or natural water fluoride concentrations.
Helte *et al.* (2021)	Cohort study Sweden	Sub-cohort of 4,306 women aged <85 years who provided bone scan, urine sample, and diet and lifestyle questionnaire	Range 0.0 ppm to 1.0 ppm	Bone mineral density; incidence of any fracture, osteoporotic fracture, and hip fracture	Higher bone mineral density and higher incidence of hip fracture were both associated with higher exposure to fluoride; no significant associations for other fractures.
Karagas *et al.* (1996)	Ecological or correlational study USA	Comparison of incidence of fractures among White men and women aged 65–89 years during 1986–1990 across geographic regions with varying fluoridation status	Counties considered fluoridated if at least two-thirds of residents had access to naturally or artificially fluoridated water at least 0.7 ppm; counties considered non-fluoridated if <10% of residents received artificially fluoridated water or where natural	Incidence of fractures of the hip, ankle, distal forearm, and proximal humerus	No association between fluoridation status and incidence of hip or ankle fracture. No association between fluoridation status and fractures of the distal forearm or proximal humerus for women; however, risk of these fractures was

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
			fluoride concentration was <0.3 ppm		higher for men in fluoridated areas.
Kurttio *et al.* (1999)	Retrospective cohort study Finland	Comparison of incidence of hip fractures in rural populations aged 50–80 years, across geographic areas with varying levels of natural fluoride in groundwater	Fluoride concentration in rural well water, median 0.1 ppm, 403maximum 2.4 ppm (analysis is stratified up to 1.5 ppm)	Hip fracture discharge rate	In the analysis of fluoride concentrations up to 1.5ppm, no clear dose-response between fluoride concentration and hip fractures was observed.
Levy *et al.* (2014)	Prospective cohort study USA	Iowa Fluoride Study: longitudinal birth cohort study (424 participants) of association between children's fluoride intake and bone mineral density outcomes at age 15 years	Total fluoride intake measured by questionnaire; most children living in optimally fluoridated areas	Bone mineral density	No associations observed between daily fluoride intake and bone mineral density.
Levy *et al.* (2018)	Prospective cohort study USA	Iowa Fluoride Study: longitudinal birth cohort study (424 participants) of association between children's fluoride intake and bone mineral density outcomes at age 11 years	Total fluoride intake measured by questionnaire; most children living in optimally fluoridated areas	Bone mineral density	Generally no or weak associations between daily fluoride intake or cumulative daily fluoride intake and bone mineral density.
Näsman *et al.* (2013)	Cohort study Sweden	Investigated associations between fluoride exposure from community water supplies and incidence	Exposure stratified as very low (<0.3 ppm), low (0.3–0.69 ppm), medium (0.7–1.49 ppm) and high (>1.5 ppm)	Incidence of hip fracture or low-trauma osteoporotic hip fracture	No association between fluoride exposure and incidence of hip fracture or low-trauma

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
		of hip fractures (60,773 fractures) in cohort aged >70 years			osteoporotic hip fracture.
Oweis *et al.* (2018)	Prospective cohort study USA	Iowa Fluoride Study: longitudinal birth cohort study (380 participants) of association between children's fluoride intake and bone mineral density outcomes at age 17 years	Total fluoride intake measured by questionnaire; most children living in optimally fluoridated areas	Bone mineral density	Generally no or weak associations between daily fluoride intake or cumulative daily fluoride intake and bone mineral density.
Phipps *et al.* (1998)	Cross-sectional survey USA	Investigated associations between natural fluoride concentrations in rural water systems and bone mineral density in 670 adults aged 60 years or older	Three rural communities with natural fluoride concentrations of 0.03 ppm, 0.7 ppm, and 2.5 ppm respectively	Bone mineral density	No differences in bone mineral density of lumbar spine, proximal femur, or forearm between low and medium fluoride communities for men or women.
Neuropsychological outcomes					
Bashash *et al.* (2017)	Prospective cohort study Mexico	Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) project: longitudinal birth cohort study (299 mother–child pairs) of association between maternal urinary fluoride in pregnancy and measures of	Natural fluoride in Mexico City ranging 0.15–1.38 ppm; fluoridated salt	Age 4: McCarthy Scales of Children's Abilities; Age 6–12: Wechsler Abbreviated Scale of Intelligence	Higher prenatal exposure to fluoride was associated with lower general cognitive scores in children at age 4 years and with lower FSIQ scores at 6–12 years old. Sensitivity analysis indicated that associations at 6–12

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
		offspring cognitive performance at age 4 and 6–12 years.			years may be limited to higher exposures.
Cancer					
Gelberg *et al.* (1995)	Case-control study USA	171 cases of osteosarcoma diagnosed 1978–1988, pair-matched to control subjects by year of birth and sex; lifetime exposure to fluoride calculated based on self-reported use of fluoride supplements, dental treatments, and drinking water and breast milk	Fluoride concentration assumed to be 1.0 ppm and 0.0 ppm in fluoridated and non-fluoridated areas respectively	Osteosarcoma risk	No association between lifetime exposure to fluoride in drinking water and osteosarcoma risk.
Levy *et al.* (2012)	Ecological or correlational study USA	Comparison of incidence of osteosarcoma among children and adolescents between states with high and low percentages of the population accessing artificially fluoridated water	States categorised as low (<30%) or high (>85%) according to the percentage of population receiving artificially fluoridated water	Incidence of osteosarcoma	No association between fluoridation status and osteosarcoma incidence during childhood and adolescence.
Schwartz (2014)	Ecological or correlational study USA	Investigated association between age-adjusted eye cancer incidence and percentage of	Percentage of population receiving fluoridated water at state level	Incidence of eye cancer	States with greater access to fluoridated water had lower incidence of eye cancer.

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
		population receiving fluoridated water			
Takahashi *et al.* (2001)	Ecological or correlational study USA	Investigated association between cancer incidences and percentage of population receiving fluoridated water 1978–1992	Fluoridation index: ratio of inhabitants receiving optimally fluoridated water to the total population of the community at state level	Incidence of cancers at 36 sites of the body	Of the 36 sites, two-thirds showed positive associations with fluoridation, one-quarter no association, and one-tenth negative associations. The authors indicate that the consistency of an association between fluoride and cancer was not adequately confirmed.
Yiamouyiannis (1993)	Ecological or correlational study USA	Compared incidence of bone cancer and osteosarcoma deaths across 26 areas with and without fluoridated community water supplies, 1973–1987	Artificially and naturally fluoridated and non-fluoridated areas (artificial fluoridation concentration 0.7–1.2 ppm)	Incidence of bone cancer and osteosarcoma among white men and women	Males under age 20 years at increased risk of bone cancer and osteosarcoma in fluoridated areas.
Endocrine conditions					
Fluegge (2016)	Ecological or correlational study USA	Investigated associations between added and natural fluoride and diabetes incidence and prevalence in 22 US states in 2005 and 2010	Optimal fluoride level defined as 0.7–1.2 ppm	Diabetes incidence and prevalence	Fluoride artificially added to drinking water to achieve optimal levels was associated with higher incidence and prevalence of diabetes. Of the three fluoridation chemicals

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
					examined as part of the study (sodium fluoride, fluorosilicic acid, or sodium fluorosilicate), fluorosilicic acid was associated with lowest incidence and prevalence of diabetes.
Liu *et al.* (2019)	Prospective cohort study Mexico	Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) project: longitudinal birth cohort study (333 participants aged 10–17 years) of association between urinary fluoride and pubertal indicators	Natural fluoride in Mexico City ranging 0.15–1.38 ppm; fluoridated salt	Tanner staging of pubertal development	Higher urinary fluoride associated with later pubertal development (pubic hair growth and genital development) for boys but not girls.
Liu *et al.* (2020)	Prospective cohort study Mexico	Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) project: longitudinal birth cohort study (536 participants aged 10–18 years) of association between urinary fluoride and cardiometabolic outcomes	Natural fluoride in Mexico City ranging 0.15–1.38 ppm; fluoridated salt	Assay of adiposity and cardiometabolic measures, including BMI; waist circumference; serum concentrations of glucose, insulin, and lipids; cardiometabolic risk score	Higher fluoride exposure associated with increased levels of cardiometabolic risk factors for adolescent girls but not boys.

Paper	Study design Country	Brief description	Exposure	Outcome	Summary of findings
Renal/hepatic outcomes					
Malin *et al.* (2019)	Cross-sectional survey USA	Investigated associations between plasma and tap water fluoride concentrations and serum indicators of kidney and liver function among adolescents	Mean fluoride concentration in tap water 0.48 ppm; 75th–95th percentile ranged 0.71–1.00 ppm	Kidney and liver parameters measured by serum markers	Higher plasma fluoride was associated with lower estimated glomerular filtration rate, higher serum uric acid concentration, and lower blood urea nitrogen concentration. Higher fluoride concentrations in water were associated with lower blood urea nitrogen concentration.
Cardiovascular outcomes					
Näsman *et al.* (2016)	Cohort study Sweden	Investigated associations between fluoride exposure from community water supplies and incidence of myocardial infarction	Exposure stratified as very low (<0.3 ppm), low (0.3–0.7 ppm), medium (0.7–1.5 ppm) and high (>1.5 ppm)	Incidence of myocardial infarction	No association between fluoride exposure and incidence of myocardial infarction

References for studies of natural fluoride

Bashash M, Thomas D, Hu H, Angeles Martinez-Mier E, Sanchez BN, Basu N, Peterson KE, Ettinger AS, Wright R, Zhang Z, Liu Y. Prenatal fluoride exposure and cognitive outcomes in children at 4 and 6–12 years of age in Mexico. Environmental health perspectives. 2017 Sep 19;125(9):097017.

Cooper C, Wickham C, Lacey RF, Barker DJ. Water fluoride concentration and fracture of the proximal femur. Journal of Epidemiology & Community Health. 1990 Mar 1;44(1):17-9.

Fluegge K. Community water fluoridation predicts increase in age-adjusted incidence and prevalence of diabetes in 22 states from 2005 and 2010. J Water Health 2016;14:864–77. doi:10.2166/wh.2016.012

Gelberg KH, Fitzgerald EF, Hwang SA, Dubrow R. Fluoride exposure and childhood osteosarcoma: a case-control study. American Journal of Public Health. 1995 Dec;85(12):1678-83.

Helte E, Donat Vargas C, Kippler M, Wolk A, Michaëlsson K, Åkesson A. Fluoride in Drinking Water, Diet, and Urine in Relation to Bone Mineral Density and Fracture Incidence in Postmenopausal Women. Environmental health perspectives. 2021 Apr 6;129(4):047005.

Karagas MR, Baron JA, Barrett JA, Jacobsen SJ. Patterns of fracture among the United States elderly: geographic and fluoride effects. Annals of epidemiology. 1996 May 1;6(3):209-16.

Kurttio P, Gusta vs. son N, Vartiainen T, Pekkanen J. Exposure to natural fluoride in well water and hip fracture: a cohort analysis in Finland. American journal of epidemiology. 1999 Oct 15;150(8):817-24.

Levy M, Leclerc BS. Fluoride in drinking water and osteosarcoma incidence rates in the continental United States among children and adolescents. Cancer epidemiology. 2012 Apr 1;36(2):e83-8.

Levy SM, Eichenberger-Gilmore JM, Warren JJ, Kavand G, Letuchy E, Broffitt B, et al. Associations of fluoride intake with children's cortical bone mineral and strength measures at age 11. Journal of public health dentistry. 2018 Sep;78(4):352-9.

Levy SM, Warren JJ, Phipps K, Letuchy E, Broffitt B, Eichenberger-Gilmore J, et al. Effects of life-long fluoride intake on bone measures of adolescents: a prospective cohort study. Journal of dental research. 2014 Apr;93(4):353-9.

Liu Y, Téllez-Rojo M, Hu H, Sánchez BN, Martinez-Mier E, Basu N, Mercado-García A, Solano-González M, Peterson KE. Fluoride exposure and pubertal development in children living in Mexico City. Environmental Health. 2019 Dec;18(1):1-8.

Liu Y, Téllez-Rojo M, Sánchez BN, Ettinger AS, Osorio-Yáñez C, Solano M, Hu H, Peterson KE. Association between fluoride exposure and cardiometabolic risk in peripubertal Mexican children. Environment international. 2020 Jan 1;134:105302.

Malin AJ, Lesseur C, Busgang SA, Curtin P, Wright RO, Sanders AP. Fluoride exposure and kidney and liver function among adolescents in the United States: NHANES, 2013–2016. Environment international. 2019 Nov 1;132:105012.

Näsman P, Ekstrand J, Granath F, Ekbom A, Fored CM. Estimated drinking water fluoride exposure and risk of hip fracture: a cohort study. Journal of dental research. 2013 Nov;92(11):1029-34.

Näsman P, Granath F, Ekstrand J, Ekbom A, Sandborgh-Englund G, Fored CM. Natural fluoride in drinking water and myocardial infarction: a cohort study in Sweden. Science of the Total Environment. 2016 Aug 15;562:305-11.

Oweis RR, Levy SM, Eichenberger-Gilmore JM, Warren JJ, Burns TL, Janz KF, et al. Fluoride intake and cortical and trabecular bone characteristics in adolescents at age 17: A prospective cohort study. Community dentistry and oral epidemiology. 2018 Dec;46(6):527-34.

Phipps KR, Orwoll ES, Bevan L. The association between water-borne fluoride and bone mineral density in older adults. Journal of dental research. 1998 Sep;77(9):1739-48.

Schwartz GG. Eye cancer incidence in US states and access to fluoridated water. Cancer Epidemiology and Prevention Biomarkers. 2014 Sep 1;23(9):1707-11.

Takahashi K, Akiniwa K, Narita K. Regression analysis of cancer incidence rates and water fluoride in the USA based on ICAR/IARC (WHO) data (1978–1992). International agency for research in cancer. J Epidemial. 2001;11:170-9.

Yiamouyiannis JA. Fluoridation and cancer: The biology and epidemiology of bone and oral cancer related to fluoridation. Fluoride 1993;26(2):83-96.