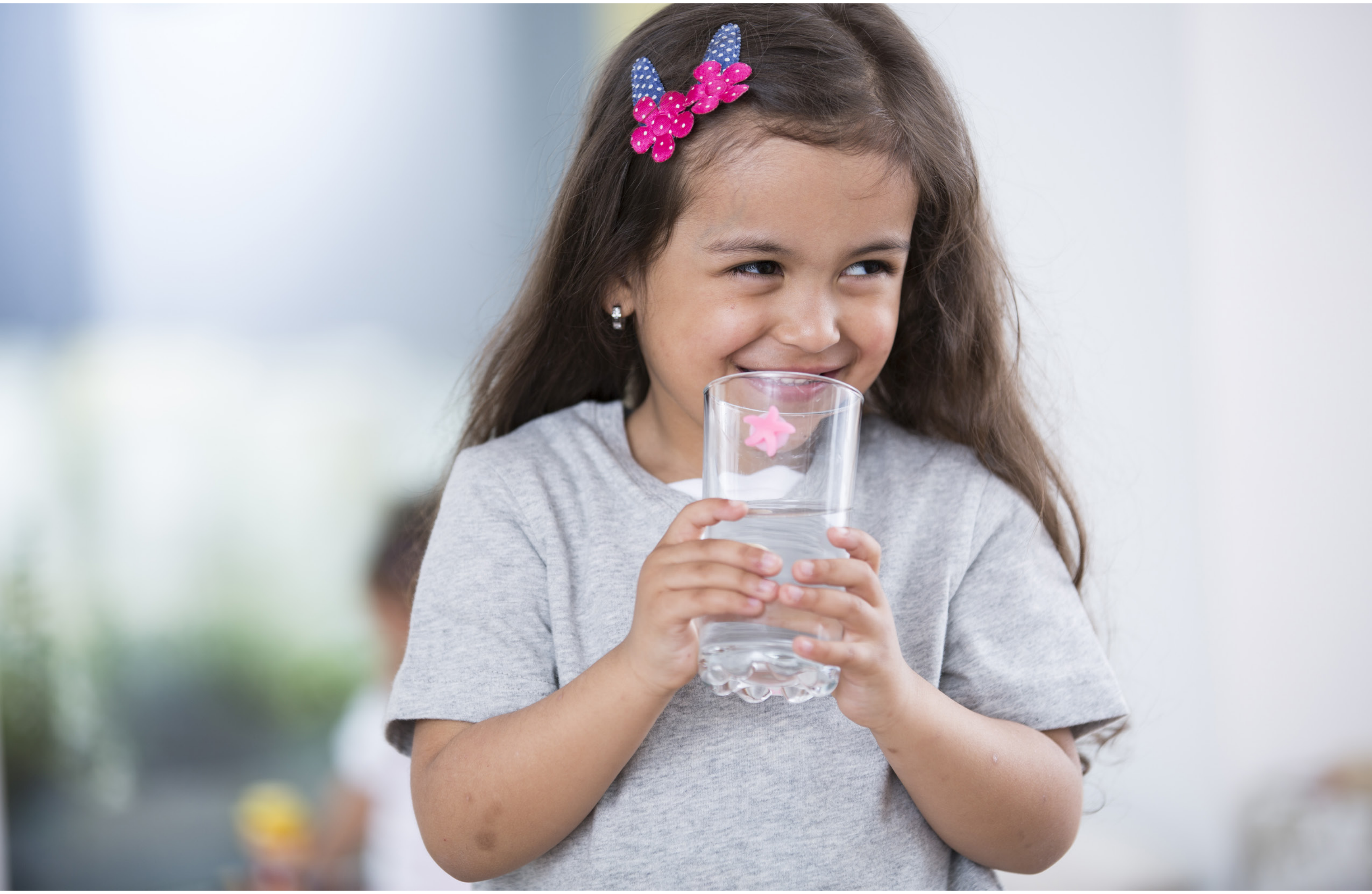
How to Read a Study about Fluoride or Fluoridation



PRODUCED BY THE BRITISH FLUORIDATION SOCIETY
AND THE AMERICAN FLUORIDATION SOCIETY

How to Read a Study about Fluoride or Fluoridation

A guide for the non-expert

Every week, new studies about health or medicine are released. Even well-educated people can struggle to understand the conclusions of these studies and how they were conducted.

When the issue of water fluoridation arises, it is often accompanied by claims for and against, with supporting evidence. It can be difficult for the layperson to decide whether evidence is legitimate, relevant and accurate, as well as whether it is used correctly in any partisan arguments. The obvious answer to the dilemma is to seek expert guidance. There is plenty of that but it is often difficult to find; unfortunately, there is also bogus "expert" advice, which is much more accessible.

This guide is intended to help those interested in the truth sort the sheep from the goats when considering the research on fluoridation. The guide is no substitute for expert opinion but is better than relying on claims being bandied about on social media or other forums.

Why would anyone bother producing bogus "expert" advice (or fake evidence) in relation to an issue such as water fluoridation? The simple answer is that if they are opposed to fluoridation the use of flawed "expert" advice can muddy the waters around efficacy and safety. And this is not a fanciful idea. For many years, the tobacco industry polluted the evidence around smoking and health to prolong the survival of cigarette smoking in the general population. They did this using various tricks (political lobbying, funding bogus/partisan research, paying for public relations expertise, and so on). So don't be surprised if some of these tactics are used by other campaigners, including those opposed to vaccines and fluoridation.

Even well-meaning people and organisations can produce flawed research. This document should help you stay close to the legitimate science.

A research service that analyses new studies for health care providers performed an assessment of about 50,000 studies. Incredibly, this assessment found that only 6% were well-designed and relevant enough to inform patient care. This is only one reason why both health journalists and their readers should be skeptical readers of new studies.

On the one hand, new studies about fluoride can offer important insights. On the other hand, new research (and how it is reported) needs to be placed in a larger context and held to key standards. Thousands of studies have been published about fluoride or fluoridation. This document offers tips for how journalists and the public can better understand scientific research about fluoride.

Contents

This document addresses a number of common errors found in published academic papers and offers 9 suggestions for assessing a research paper:

This guide describes these analysis steps with examples from actual published papers. The Appendix contains a spoof research paper with built-in flaws, for readers to test their analysis skills.

1 Read beyond the abstract.

The abstract is intended to summarize a study, but sometimes it may omit a critical fact or detail. For this reason, one cannot rely on an abstract to fully understand a study's conclusions or its potential weaknesses. For example, the abstract for this fluoride-related study reported this conclusion:

> "Exposure to increasing levels of fluoride in tap water was associated with non-verbal intellectual abilities; the effect was more pronounced among formula-fed children."

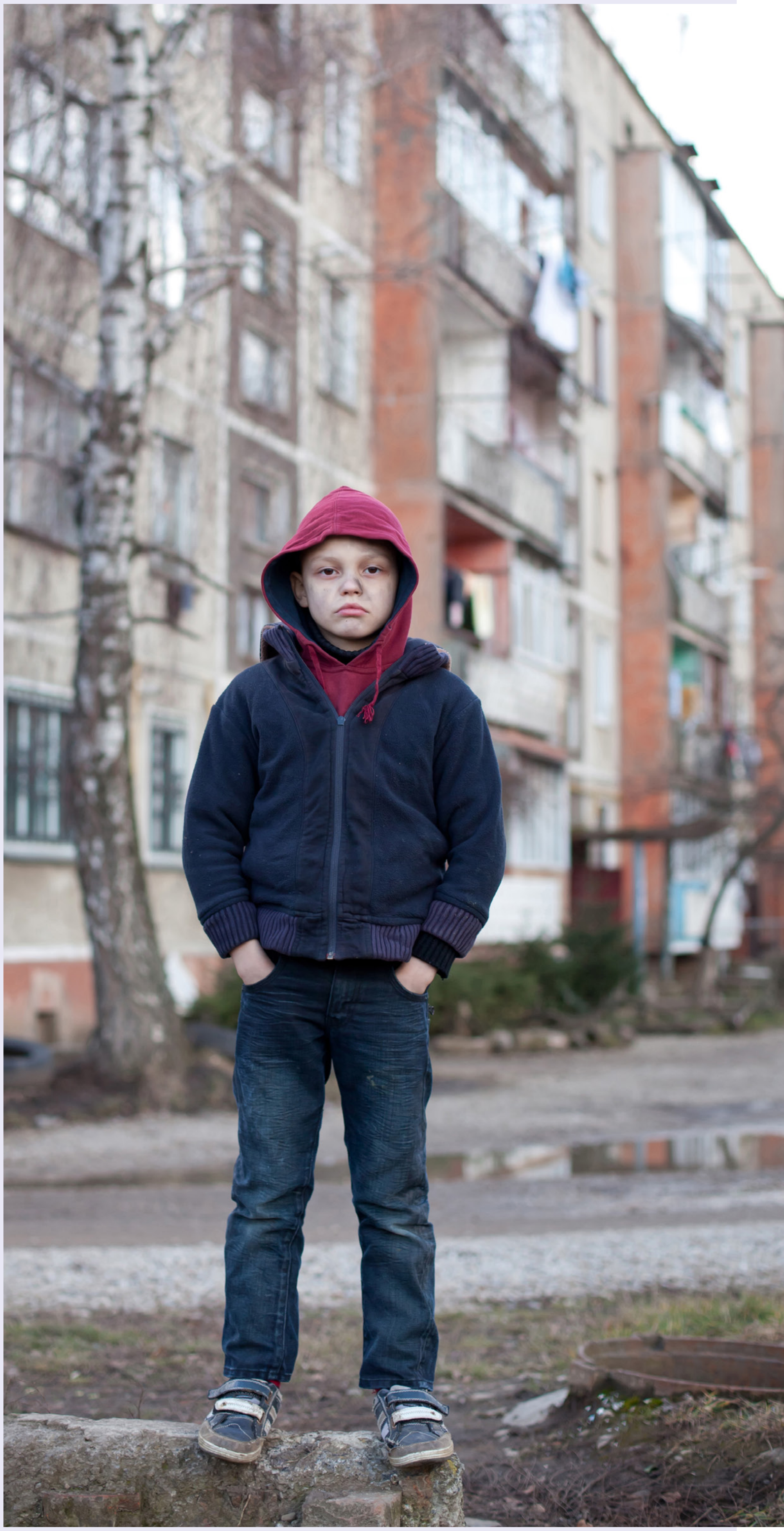
The conclusion did not include any cautionary phrases. Basically, the authors told us they found a link between fluoride in water and something harmful. Yet a very important sentence in this paper could be overlooked because it was buried in the middle of this study. Here is that sentence:

> "The association between water fluoride concentration and [full-scale IQ] must be interpreted with caution, however, because the association became non-significant when two outliers were removed."

This statement raises a red flag. If the IQ scores of only two children could skew the results of a study, the authors have an obligation to write their conclusions with much greater caution and specificity.

The lesson is clear: abstracts can be helpful, but they sometimes tell only part of the story. It is critical to carefully read an entire study to fully grasp its meaning and relevance.

2 Consider the confounding factors.



A confounding factor (or confounder) is something different from the main topic of a study that, nonetheless, could affect the study's results. Since most health outcomes have multiple contributing causes, there are many potential confounders that should be considered by researchers.

Suppose researchers want to study whether adolescents who live close to several lakes in three adjoining counties are more likely to drown than those who live far away from the lakes. The researchers might gather data from recent years about people who have drowned in these lakes. They might learn that most adolescents who drowned in one of these lakes had indeed been living within five miles of a lake. Proximity is an obvious risk factor (and subject to methodological distortion), but other factors may be more important and amenable to change. If this research paper identified "living near a lake" as a strong risk factor for drowning without acknowledging other factors, this could mislead the public by ignoring confounders that might have an equal or higher association with a person's risk of drowning.

A broader analysis by these researchers probably would have shown that people who had been swimming or boating after sunset were more likely to be drowning victims. People who were swimming alone or who had consumed alcoholic beverages right beforehand might also have a higher risk of drowning. Without considering confounders like these, the researchers could present a conclusion that overstates the proximity to a lake as a risk factor.

Confounders should be a major consideration in research about fluoride's effects. A number of studies exploring a link between fluoride and children's IQ scores have been published. None of them (to our knowledge) has taken into account the IQ scores of mothers. This is noteworthy because years of research have shown that maternal IQs influence children's scores. The Green-Till study (2019) is one example of a research paper for which no data on maternal IQs was available to consider.

There were other confounders that the Green-Till paper and other fluoride studies either could not or did not account for. After reading the Green-Till study, a British expert in statistical analysis wrote that "nothing was taken into account that occurred between birth and the age of 3 or 4 when the child's IQ was measured." One factor that was *not* considered was the children's lead exposure during early childhood. Although the Green-Till authors had access to maternal lead exposure, they had no data about lead exposure that the children experienced during the three to four years before their IQs were tested.

Moreover, another potential confounder was identified by the 31 toxicologists who reviewed the Green-Till paper in an article published by the *Archives of Toxicology*. As these experts explained, the IQ tests were performed "only once between the age of 3 and 4 years, but the exact age of the children at the time point of the test has not been considered in the statistical analysis. This may be problematic, because the IQ of children changes strongly between 3 and 4 years." In other words, the findings could have been skewed if, compared to children in the fluoridated cities, the children in the non-fluoridated cities were more likely to have been tested for IQ at age 4, instead of at age 3.

It can be difficult to account for every potential confounder, but the authors of a study should strive to consider as many as possible. If a lack of data leaves several key confounders unaccounted for, the authors should write their conclusions with an appropriate degree of caution.

3 Beware of how fluoride is measured.

The strength of a fluoride study is significantly based on the exposures that were measured, as well as how they were measured.

First, the methodology should be transparent, providing clear and complete information on the concentration of fluoride that was studied. Opponents of fluoridation often circulate a 2012 research paper that reviewed 27 studies, nearly all of which were conducted in China. Yet only 17 of these studies disclosed:

- the fluoride levels that were evaluated
- the ages of the children who were studied

When basic information like this is missing from a study, it should raise a red flag to journalists and the public.

Second, it is important to know the method used to measure fluoride exposure. Some methods are scientifically more reliable than others. Several studies have measured fluoride through urine samples. A number of public health experts have questioned the reliability of urine samples for measuring fluoride exposure.

An expert in environmental toxicology has questioned the reliability of measuring exposures this way. Alastair Hay, Professor (Emeritus) of Environmental Toxicology at the University of Leeds, has observed that because fluoride has a short half-life, "urine concentrations vary hugely".

The authors of a 2011 research paper wrote that "it can be concluded that, at this time, urinary fluoride excretion has a very limited value as a biomarker of individual fluoride exposure."

In addition, not all urine samples are equally reliable. A spot urine sample means only one sample is taken, preferably early in the morning. Studies that use a 24-hour urine sample have collected all the urine produced by study subjects over the course of a 24-hour period or day. The widely publicised Green-Till paper (2019) based its analysis on so-called "spot samples" of urine to measure women's fluoride exposure during pregnancy. But experts stress that 24-hour urine samples are more reliable ways to measure exposure.

4 Look for any inconsistencies.

A 2016 research paper was given a scary headline that linked water fluoridation with diabetes in the United States, but it was very misleading. In fact, the paper itself later explained why this headline is not accurate.

There are three fluoride additives that are used for the process of water fluoridation. The author of the 2016 paper found that one particular additive (sodium fluoride) was significantly associated with a higher incidence and prevalence of diabetes. If that were the end of the story, the paper's headline might make sense. But hold on.

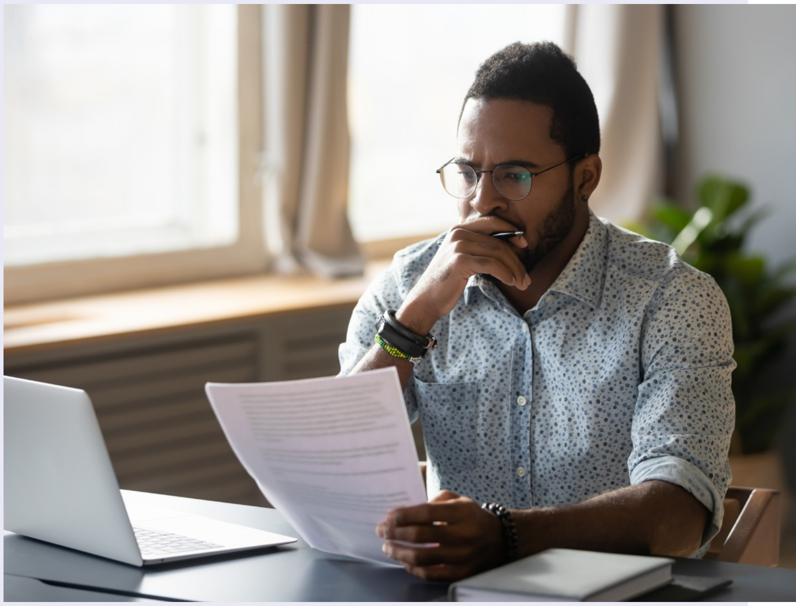
The other two fluoride additives that are used for water fluoridation are:

- fluorosilicic acid
- sodium fluorosilicate

The author reported that these two additives had "significantly negative associations" with diabetes. In other words, these two fluoride additives were linked to a *lower* incidence and prevalence of diabetes. It's worth noting that most water systems in the United States use one of these two additives to fluoridate their drinking water.

This is a glaring inconsistency. Regardless of what type of fluoride additive is used for fluoridation, water companies that engage in fluoridation strive to maintain fluoride at the same optimal level. If, indeed, the presence of fluoridated water could predict a county's rate of diabetes, then the association should be consistent regardless of which fluoride additive is used. This inconsistency demonstrates why this study's headline misrepresents the actual results.

A different inconsistency exists among the Canadian cities involved in the Green-Till study (2019). Although the authors assert that fluoride is associated with lower intelligence, the IQ scores in each of these cities do not show a clear pattern. The master's thesis on which the study was based showed (Table 3) that the full-scale IQ score in the non-fluoridated city of Kingston was 8.05 points lower than the baseline score of non-fluoridated Vancouver. Yet there was virtually no difference in IQ scores between fluoridated Hamilton and non-fluoridated Montreal.



5 Consider the relevance.

Thousands of studies have been published that examine the impact of fluoride exposure, and many of these papers have assessed fluoride exposures in industrial settings. Because of the nature of this exposure, it has little or no relevance to the kind of exposure that occurs in a fluoridated community.

Relevance is also an issue when studies are based on fluoride exposures that are far higher than those normally observed in North America and Western Europe.

Critics of fluoridation have disseminated a 2012 research paper that reviewed a group of studies that were mostly performed in China. In all, 17 studies provided complete data on the fluoride concentrations in water and the ages of the children whose IQs were tested. However, most of these 17 studies tested naturally occurring fluoride levels that were more than double the level used for water fluoridation in the United States and United Kingdom. This casts serious doubt on the relevance of these studies.

In addition, for these 17 studies, the average fluoride concentration in the low fluoride "reference" samples of water – which were associated with **higher IQ scores** – was 0.71 milligrams per liter. This is close to the recommended concentration for water fluoridation.

Generally, many studies exploring the association between fluoride and neurotoxicity are based on fluoride levels that are much higher than those used for water fluoridation. This is one reason why the National Academies of Sciences, Engineering and Medicine instructed a committee of researchers in 2021 to "make it clear" that their draft paper about fluoride and neurotoxicity "cannot be used to draw any conclusions" about the fluoride exposures that are "typically associated with drinking-water fluoridation".

6 Explore the quality or reputation of the journal.

The journal in which a study is published is a relevant issue to consider. Not all journals have a reputation for conducting peer review in a rigorous manner. A prime example of this is *Fluoride*, a journal that is frequently cited by opponents of fluoridation.

Fluoride has a reputation of having published "a number of flawed or scientifically incomplete studies." For example, a 2005 article in *Fluoride* claimed to have linked fluoridation with crime. The author, having no apparent scientific credentials in the fields of criminology or epidemiology, explained his methodology:

> "The (crime) stories were selected, based on their content and on my intuition, from my routine daily reading, rather than from a methodical or exhaustive search using, for example, keywords or search engines."

From 1982 until his death in 2013, Albert Burgstahler served as editor, co-editor or science editor of *Fluoride*. Burgstahler, a chemistry professor at the University of Kansas, made no secret of his stern opposition to water fluoridation. The editor-in-chief of *Fluoride* is Dr Bruce Spittle, a psychiatrist in Dunedin, New Zealand, whose views are stridently anti-fluoridation, as demonstrated by the title of a book he authored: "Fluoride Poisoning." As of August 2021, Spittle remained editor-in-chief of *Fluoride* more than five years after he was convicted of sexual assault of a patient and more than three years after his registration as a health practitioner was cancelled by a medical tribunal.

A related concern is the presence of predatory journals. These journals often require a publication fee, may not perform a robust peer review process, and/or may not conduct checks related to plagiarism or other ethical issues. Many researchers have been misled into submitting papers to such journals. A study of 46,000 European researchers revealed that 5% of them had published papers in a predatory journal.

7 Review the study's references.

There are several types of bias that can affect a research paper. One of them is confirmation bias, which occurs when researchers – consciously or unconsciously – look for findings or make interpretations that confirm the ideas or opinions they already hold. The references that researchers include at the end of their paper sometimes offer signs of confirmation bias.

- Did the authors cite journal articles or reports from reputable and scientifically rigorous sources? Or did they cite articles from sources that lack scientific rigor or might be ideologically driven?
- Did the authors overlook articles that came to different conclusions?

A number of fluoride research articles include references to questionable sources. For example, a 2020 fluoride-related paper cites several obscure or irrelevant papers, including an article published in a journal that is managed by the Institute of the Chinese Silkworm. The Green-Till study includes several references from *Fluoride*, whose reputation was addressed previously in this document.

In addition, one of the Green-Till references was a research paper co-authored by Declan Waugh. The concerns raised in the Waugh paper were addressed and dismissed two years later by the Food Safety Authority of Ireland, which issued a report stating its conclusion that "there is currently no scientific basis for concerns about the safety of children and adults" from exposure to fluoride from foods and beverages". Public health experts, journalists and others have noted Waugh's record of misrepresenting reports and health data, as well as promoting a bizarre theory connecting fluoride and sexual orientation. Nonetheless, some fluoride researchers continue to cite Waugh's papers.

8 Look for the sample, its size, its drop-out rate, and how representative it is.

Although there is no “magic number” for an appropriate study sample, **larger samples generally strengthen the confidence that a study’s results are valid** and are not due to sampling error.

In 2012, opponents of fluoridation widely disseminated a research article that reviewed a cluster of 27 studies about fluoride. A significant share of these studies (44%) had sample sizes of fewer than 200 participants each.

The website of the Fluoride Action Network, an anti-fluoride group posts a 1983 study that links fluoride and dermatitis. Many visitors to the site might miss a key detail of the study. Its sample was very small: only 21 people. The infamous 1998 research paper by Dr Andrew Wakefield linking autism to measles, mumps and rubella vaccines (later retracted) raised skepticism partly because its sample size was only 12 children.

Dropout rates are important too. Sometimes, participants who begin a study fail to show up for tests or do not respond to follow-up requests for information. For this reason, some attrition is to be expected and does not reflect poorly on the researchers. But some studies have relatively high drop-out rates. Regardless of why people drop out of a study, this decline in participation is worth noting because it can lessen confidence in the researchers’ findings. The 1983 study cited earlier started with 21 participants, but one-third of the participants dropped out before the research could be completed. This left a very small sample.

Finally, it’s worth considering whether a study’s sample was representative of the larger population. If it was not, a study’s findings may have little or no relevance. The Green-Till study (2019) relied on the Maternal-Infant Research on Environmental Chemicals (MIREC) data cohort. The women who agreed to participate in MIREC were older, much less likely to smoke (5.9% vs. 10.5%), and much more likely to have attained a university-level degree (62.3% vs. 35.1%) than pregnant women overall. Researchers who summarised the MIREC cohort noted that its data “may not be generalisable to the Canadian population or to each of the recruitment sites as the study is not population-based.”

9 Consider how the analysis was conducted.

The analysis of a study should be consistent with its sampling design. Samples may be recruited randomly from a list of telephone numbers or a group/cluster of people may be selected such as schoolchildren. Most fluoride-IQ studies have sampled clusters of children or pregnant women. Children at the same school in the same town tend to have similar demographic, health and social characteristics. Although 82 children may attend a primary school in a Chinese village, the realities of "clustering" mean that researchers should view the school as a sample of one, instead of a sample of 82. In other words, in most instances, the number of clusters matters more than the number of children.

For these reasons, a study that draws its sample from clusters (e.g., 10 or 100 primary schools) should conduct its analysis differently than studies whose samples are drawn randomly from a diverse geographic area. Otherwise, comparing mean IQ between higher and lower fluoride areas might show a difference when none exists in the underlying population from which the sample was drawn.

When cluster sampling is used, the standard error will be higher. To avoid this effect, researchers should account for clustering in their analysis, for example, by calculating the mean IQ and the standard error differently than they would with a simple random sample. Unfortunately, a number of fluoride-IQ studies have not accounted for clustering in their analysis to avoid potential distortion of the effect.

The U.S. National Academies of Sciences, Engineering and Medicine (NASEM) raised this issue in a 2021 research review. In its peer review of a draft monograph about fluoride, NASEM urged a research committee to "examine the studies included in [its] meta-analysis in greater depth to determine whether each study properly accounted for its design because not doing so could invalidate the meta-analysis results".

Appendix

❶ Read beyond the abstract.
❷ Consider the confounding factors.
❸ Beware of how fluoride is measured.
❹ Look for any internal or external inconsistencies.
❺ Consider the relevance.
❻ Explore the quality or reputation of the journal.
❼ Review the study's references.
❽ Look for the sample size and dropout rate.
❾ Consider how the analysis was conducted.

A spoof research paper is available for you to scrutinize. See if you can spot the faults using the steps outlined in this guide.