APPENDIX 6-A. Kim review (includes discussion of Bassin 2006)

Chris Neurath

Two case-control studies of fluoride and osteosarcoma
Review of Kim et al. (2011) study; comparisons to Bassin et al. (2006)

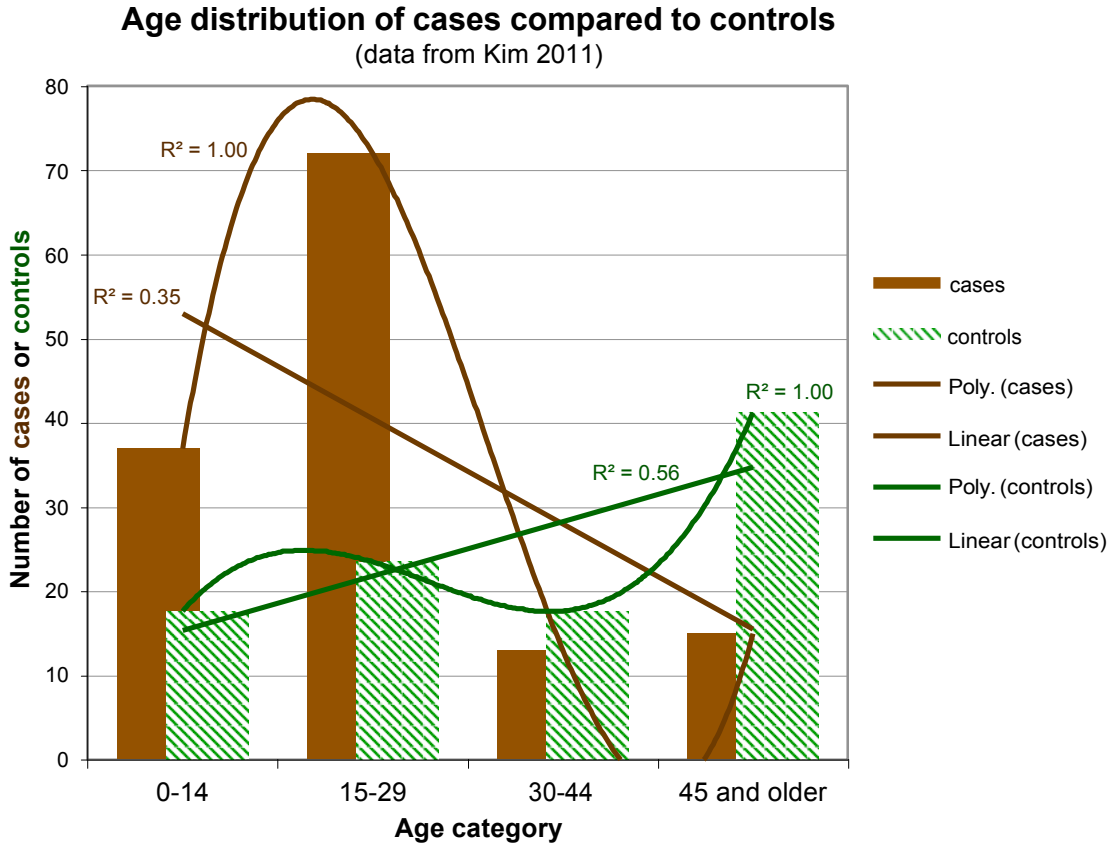
A study by Bassin et al. [2006] found a strong link between fluoride (F) exposure and osteosarcoma. A more recent study by Kim et al. [2011] failed to find an effect. The two studies were part of a single larger overall project, but used two different data sets and methods. Only one other substantial case-control study of fluoride and osteosarcoma exists [Gelberg 1995], so the Bassin and Kim studies deserve careful scrutiny. Issues of confounding, selection bias, criteria for controls, and statistical power, may explain the different outcomes.

Both papers analyzed data from case-control subject sets recruited from metropolitan tertiary care hospital orthopedics departments, but at different time periods: 1989-1992 for Bassin, 1993-2000 for Kim. Kim used bone F as the measure of exposure, while Bassin used lifetime history of F exposures from water, F toothpaste and F supplements to estimate age-specific exposure. Bassin matched on age, gender, and distance from hospital, while only a small percentage of Kim's subjects were matched on any variables. Six key differences in the studies may explain the different results.

1. **Confounding by age.**
2. **Kim's study had low statistical power.**
3. **Kim combined sexes, reducing specificity.**
4. **Kim used no exposure-timing information.**
5. **Selection bias due to controls' diseases potentially being affected by fluoride.**
6. **Possible confounding by geographic location.**

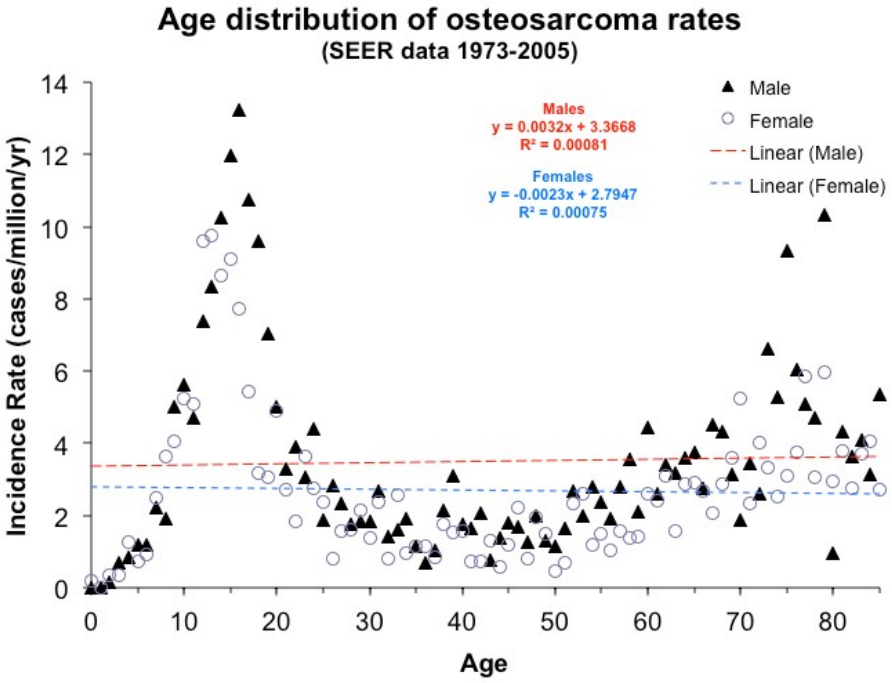
1. Confounding by age. Kim's controls had a much older age distribution than cases (median age 41.3 versus 17.6), yet had similar bone fluoride levels. Kim reported a moderately strong positive association between age and bone fluoride which is consistent with other studies [Jackson 1958, Zipkin 1958, Eble 1992, Charen 1979]. Therefore, daily average fluoride exposures must have been higher in Kim's cases than controls, roughly twice as high based on the commonly found linear age to bone fluoride relationship. The cases were less than half as old yet during their shorter lifetimes they accumulated as much bone fluoride as the older controls. The age distribution of cases is seen to be much different from controls in Figure 1. Linear regression fitting is poor while polynomial of degree 3 fitting is good.

Figure 1. Age distribution of cases compared to controls. With comparison of fitted regression functions: linear and polynomial of degree 3.



For Kim to avoid strong bias from age confounding, proper age adjustment is required. The relationship between age and osteosarcoma risk is far from linear. Figure 2 shows that osteosarcoma in the general population has a very sharp peak incidence rate at age 15, then rapidly declines to a low level until a gradual rise begins around age 50. The causes of osteosarcoma for older adults are believed to be different than for those under age 25 [Troisi 2006, Savage 2011]. The shape of the curve for under-25 year olds is an inverted-U-shape, with a very sharp peak. The sharpness of this peak is greater than for any other type of cancer. Despite this dramatically non-linear relationship between age and osteosarcoma risk, Kim apparently applied no transformations to the age variable in her multivariable analyses, so that age was treated as having a linear relationship to osteosarcoma risk. Age did not meet the linearity assumption for multivariable regressions. Applying a linear relationship to model the sharp inverted-U-shape relationship between age and osteosarcoma risk produces a flat regression line with a zero correlation coefficient, as shown in Figure 2.

Figure 2. Relationship between age and osteosarcoma rates in large US cancer registry.



Therefore, Kim's method of age adjustment likely failed to provide any adjustment, allowing strong residual confounding of the relationship between bone fluoride and osteosarcoma. A description of this situation when a variable has a U-shaped relationship to the outcome, is described by Katz (2011):

> "The simplest method of incorporating an interval-independent variable that does not have a linear relationship with outcome is to transform it so that it does fulfill the linearity assumption."
>
> "At times you may find that there is a U-shaped relationship between your interval-independent variable and your outcome. For example, Figure 4.3 shows a U-shaped relationship between cholesterol level and all-cause mortality in a sample of 1102 women.[3] Mortality is highest for women with the lowest and the highest values of cholesterol. When the investigators treated cholesterol as an interval variable, there was no significant relationship between cholesterol level and mortality, because the two trends statistically cancel each other out." [Katz 2011; p. 80-81]

Just as Katz's example showed no relationship when a variable did not have a linear relationship to outcome, Kim's method of controlling for age will produce no control of age.

Since bone fluoride increases with age, and the controls were much older than cases, this bias would be in the direction of reducing the odds ratio and potentially even causing higher bone fluoride to appear protective for osteosarcoma.

With Kim's much older controls having very similar bone fluoride levels as the younger

cases, a proper interpretation of this study is that bone fluoride is a strong risk factor for osteosarcoma. Thus, the Kim study corroborates the Bassin study, rather than refutes it as many have concluded.

The size of the effect cannot be determined without re-analysis of Kim's data with proper age adjustment, but some perspective can be gained by comparing Kim's fluoride exposures to those in Bassin's study, where large effects were found. Bassin's cases had only about 15% greater fluoride exposure at age 7 than matched controls, yet this was enough for Bassin to find a large effect. Kim's cases are estimated to have had roughly double the average annual exposure as the controls.

Kim also conducted analyses on age-matched cases and controls, but the numbers of matched cases and controls were very small, and age matching does not provide any control of confounding. In fact, it can exacerbate confounding, as described by Rothman [1998, p.151].

2. Kim's study had low statistical power. In the under 20 years old age group, Kim had about $1/10^{th}$ the number of control subjects as Bassin, likely giving numbers too small to detect effects. Bassin restricted her analyses to those under age 20 because previous studies found associations between fluoride and osteosarcoma only in this age group [Hoover 1991, Cohn 1992]. Kim acknowledged her sample size for age <20 was too small to conduct analyses. Instead, Kim limited analyses to all ages combined and to all ages <45.

3. Kim combined sexes, reducing specificity. In contrast, Bassin looked at genders separately because previous evidence suggested a stronger association between fluoride and osteosarcoma in males than females [NTP 1990, Hoover 1991, Cohn 1992]. Kim's combined genders would dilute any effect that might be greater in males. Kim did include gender in her final regression models, but does not report whether the outcome was affected by gender, nor did she report effect size by gender.

Kim also had a different male/female sex ratio for cases (1.1) compared to controls (2.4), which raises concern for selection bias.

4. Kim used no exposure-timing information. Bassin estimated fluoride exposure during each year of life from histories of residential drinking water, bottled water use, and fluoride measurements of private well water. Bassin also adjusted for use of dental fluoride supplements. Kim instead used bone fluoride to estimate total fluoride exposure. Bone fluoride concentration reflects cumulative long-term exposure but is uninformative regarding timing of exposure. The half-life of fluoride in bone is estimated at up to 20 years [NRC 2006]. Bassin found substantial differences in risk from exposures occurring at different ages. Her work is unique in looking at exposures during specific age windows; an approach which can detect elevated risk that could otherwise be obscured as shown by the US EPA and others [EPA 2005a, EPA 2005b, Rothman 1981, Murdoch 1992].

Osteosarcoma's extremely sharp incidence rate peak around age 15 suggests exposure during a narrow time window of vulnerability may be important [Bassin 2006, Troisi 2006]. Kim's use of bone fluoride, which essentially averages long-term exposure, may therefore underestimate risk.

Bassin looked at single-year exposure age windows and found exposure at ages 6-8 in males produced the greatest risk of osteosarcoma, with diagnoses peaking about 8 years later at ages 14-16. This suggests a latency period of roughly 8 years.

This implied latency period can be compared to that for a cohort of patients injected with radium-224 as a medical treatment, who later developed bone cancers. This cohort had an extremely elevated incidence of osteosarcoma, and the exposure timing was precisely known from medical records because Ra-224 has a half-life of only 3.5 days. Most of the bone sarcomas (76%) occurred in patients injected when they were under 21 years old, with a mean age of exposure around 10 years old for this group. Amongst 56 bone cancer cases, a peak wave of diagnoses occurred about 8 years after exposure. The sharpness and height of this peak is almost as great as that found for osteosarcoma in the general population. Neither dose nor age at injection seemed to affect this latency period [Nekolla 2000, Chmelevsky 1988].

Due to the sharpness of the osteosarcoma incidence peak around age 15, Troisi [2006] speculates initiation occurs during *in-utero* development. But the quantitative evidence from the Ra-224 injected cohort suggests an initiating exposure occurs not during pregnancy but in accord with Bassin's findings of highest risk during ages 6-8. Recent mathematical modeling of osteosarcoma age-incidence curves illustrates how acceleration of tumor growth during the adolescent growth spurt can sharpen the peak, even if the tumor initiations occurred over a broader time window [Valberg 2012].

Finally, Bassin's method of estimating historical fluoride exposure is noteworthy for its careful checking of information from multiple sources, rather than relying on a single source of information. Other studies of fluoride exposure have often used a single source of information, such as the CDC Fluoridation Census. Bassin found that errors, ambiguities, and missing information in such sources could lead to considerable misclassification [Bassin 2004].

5. Selection bias due to controls' diseases potentially being affected by fluoride.
Kim's controls were patients with non-osteosarcoma forms of bone cancer. While little research has been conducted on the association between fluoride and these rarer types of bone cancer, it is biologically plausible that fluoride—a bone-seeking element—could increase the risk of other types of bone cancer besides osteosarcoma. [NTP 1990, Hoover 1991, Kim 2011]. To the extent fluoride might increase the risk of non-osteosarcoma bone cancers, Kim's use of such controls would obscure an association between fluoride and osteosarcoma.

A potential weakness of Bassin's study, relative to Kim's, is a possible opposite direction selection bias due to Bassin's use of non-cancer orthopedic patient controls. A recent fluoride toxicology review [NRC 2006] found a number of effects on bone, joints, and other systems. Most were detrimental to health so could produce selection bias of orthopedic controls toward higher fluoride exposure, thus lowering Bassin's ORs, not spuriously inflating them. But for bone fractures, one study suggested an increased risk for those drinking low fluoride water (<0.3 ppm) compared to those drinking 1 ppm [Li 2001].

Assuming water with 1 ppm fluoride does confer protection against fractures, compared to lower fluoride exposures, we conducted a sensitivity analysis to see what degree of protection and what proportion of controls with fractures would be required to bias Bassin's study sufficiently to produce a spurious OR = 5.5 if in fact the true OR = 1.0 [Rothman & Greenland 1998]. If fewer than 33% of controls were fracture patients, it was impossible to create sufficient bias no matter how great the protective effect of fluoride against fractures. More than 75% of controls would need to be fracture patients and 1 ppm fluoride would need to have a strong protective effect against fractures (OR = 0.2) before it would be possible to bias the osteosarcoma OR enough to reach 5.5. These conditions seem highly unlikely. [Details of the sensitivity analysis in Appendix 6-B.]

6. Possible confounding by geographic location. Kim's statistical analysis did not attempt to adjust for distance from hospital, yet she matched some of her subjects on distance from the recruiting hospital. Distance from hospital is a potential confounder because 100% of the 9 recruitment hospitals were located in fluoridated metropolitan areas, whereas the proportion of the US population living outside urban areas who have fluoridated water is roughly 50%. Kim had almost twice as many cases who had never lived in an urban area (13.1%) compared to controls (7.8%), so her controls were more likely to have lived in fluoridated areas. Since some subjects were matched by distance from hospital, Kim's analyses should have controlled for distance to reduce bias. Confounding from this selection bias may have produced a spuriously lower OR, since the controls were not from the same underlying population of the controls, and they were selected such that they were more likely to have lived in fluoridated areas.

Estimating the lowering of Bassin's effects if her study had used Kim's methods.

As a result of the aforementioned six weaknesses, Kim's study had less ability than Bassin's to detect an association between fluoride and osteosarcoma. They may explain why Kim found only a small positive effect that did not reach statistical significance (OR = 1.3) while Bassin's age-specific approach found a statistically significant effect with an OR as large as 5.5 in the most sensitive group: males under age 20 who had been exposed at age 7. No opposite-acting weaknesses, which could have caused Bassin's study to overestimate ORs, were identified.

To illustrate the magnitude of some of Kim's weaknesses, we have estimated how three

of them, if applied to Bassin's study, would have reduced Bassin's effect size:

A. Bassin used age-specific exposures, while Kim looked at bone fluoride, which in children essentially represents lifetime cumulative exposure because the half-life of F in bone is estimated at 20 years [NRC 2006]. A rough estimate of the risk from cumulative lifetime exposure for 15 year olds in Bassin can be obtained by averaging the exposure-age-specific ORs shown in her Figure 2 [Zeiss 1992]. For males, this reduces the OR from a maximum of 5.5 (for exposure at age 7) to 3.0 for lifetime exposure. Similarly, average lifetime OR for females is 1.5.

B. Kim combined males and females, while Bassin analyzed them separately. Combining the lifetime ORs for Bassin's males and females estimated above, weighting for number of each gender, gives an OR = 2.5.

C. Bassin restricted her study to under-20 year olds, while Kim included all ages. If we assume that osteosarcomas in those ≥20 are not associated with F, as suggested by some researchers [Mirabello 2009, 2011; Savage 2011, Douglass 2006, Hoover 1991, Bassin 2006, Gelberg 1995], the ≥20 age group in Kim can be assigned an OR = 1.0. To calculate an all-ages OR, combine this OR of 1.0 for ≥20 year olds with the OR of 2.5 calculated above for Bassin's <20 year olds, weighting by the numbers of subjects in the ≥20 and <20 age categories in Kim's study, to get OR = 1.3.

Applying Kim's methods to Bassin's data yields the exact OR Kim found in her own data. This exercise shows Kim's combining of exposure periods, sexes, and ages could explain her small effect size relative to Bassin's.

Kim's inadequate age adjustment may also have played a large role in producing biased estimates that greatly underestimated the risk of osteosarcoma from fluoride.

In conclusion, the Kim study suffers limitations which could readily explain why it failed to confirm Bassin's findings. Indeed, careful examination of Kim's published data suggests it actually corroborates and strengthens Bassin's findings.

Bassin [2006] foresaw most of the limitations affecting Kim's study and called for:

"… studies with larger numbers of osteosarcoma patients, with incidence under age 20, that examine age---specific and sex---specific associations are required to confirm or refute the findings of the current study."

Kim's paper addressed none of these recommendations.

If the full data available to Kim, including residential histories for fluoride exposure, were analyzed with more robust methods, it might confirm Bassin's findings. To help settle this issue, the Harvard/NCI project should share its data with independent researchers. The project is now 20 years past its original intended completion date. It is time to release this data for others to examine from fresh viewpoints.

REFERENCES

Bassin EB, Wypij D, Davis RB, Mittleman MA. Age-specific fluoride exposure in drinking water and osteosarcoma (United States). *Cancer Causes Control*. 2006;17(4):421–428.

Charen J, Taves DR, Stamm JW, Parkins FM. Clinical Investigations Bone Fluoride Concentrations Associated with Fluoridated Drinking Water. *Calcif Tissue Int*. 1979;99(1):95–99.

Chmelevsky D, Kellerer AM, Land CE, Mays CW, Spiess H. Time and dose dependency of bone--- sarcomas in patients injected with radium---224. *Radiat Environ Biophys*. 1988;27(2):103–114. Available at: http://www.ncbi.nlm.nih.gov/pubmed/3164868.

Cohn PD (1992) A brief report on the association of drinking water fluoridation and the incidence of osteosarcoma among young males. Trenton NJ: New Jersey Department of Environment Protection and Energy and the New Jersey Department of Health.

Douglass CW, Joshipura K. Caution needed in fluoride and osteosarcoma study [letter]. *Cancer Causes Control*. 2006;17(4):481–482.

Eble DM, Deaton TG, Wilson FC, Bawden JW. Fluoride Concentrations in Human and Rat Bone. *J Public Health Dent*. 1992;52(5):288–291.

EPA Risk Assessment Forum (2005) Guidelines for Carcinogen Risk Assessment. US EPA, Washington DC. EPA/630/P-03/001B.

{Page 2-6 "In general, greater weight should be given to studies with more precise and specific exposure estimates."

Page 3-4 "Timing of exposure can also be important. When there is a susceptible lifestage, doses during the susceptible period are not equivalent to doses at other times, and they would be analyzed separately"

Page 3-26 "(Murdoch et al., 1992)." Averaging exposure over time may underestimate risk by a factor of 2 to 5.}

Environmental Protection Agency. (2005). *Guidelines for carcinogen risk assessment*. Washington DC. Retrieved from http://www.epa.gov/ttnatw01/cancer_guidelines_final_3-25-05.pdf

Environmental Protection Agency. (2005). *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens. EPA/630/R-03/003F*. Washington DC. Retrieved from http://www.epa.gov/ttnatw01/childrens_supplement_final.pdf

Gelberg KH (1994) Case-control study of childhood osteosarcoma [dissertation]. New Haven CT: Yale University.

Gelberg KH, Fitzgerald EF, Hwang SA, Dubrow R. Fluoride exposure and childhood osteosarcoma a case-control study. *Am J Public Health*. 1995;85(12):1678–1683.

Nekolla EA, Kreisheimer M, Kellerer AM, Kuse-Isingschulte M, Gössner W, Spiess H (2000) Induction of Malignant Bone Tumors in Radium-224 Patients: Risk Estimates Based on the Improved Dosimetry. Radiat Res 153: 93-103.

{"In the study cohort, 56 malignant bone tumors occurred in a temporal wave that peaked 8 years after exposure, whereas less than one case would have been expected during the follow-up." [Nekolla 2000]}

Greenland S (1998) Basic methods for sensitivity analysis and external adjustment. In: Rothman KJ, Greenland S, Modern Epidemiology. Philadelphia PA: Lippincott Williams & Wilkin, p. 343-357.

Hoover RN, Devesa SS, Cantor KP, Fraumeni JH (1991) Time trends for bone and joint cancers and osteosarcomas in the Surveillance, Epidemiology and End Results (SEER) program National Cancer Institute. Appendix F 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. Washington DC: Public Health Service, U.S. Department of Health and Human Services.

Hoover RN, Devesa SS, Cantor KP, Lubin JH, Fraumeni JH (1991) 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. Washington DC: Public Health Service, U.S. Department of Health and Human Services.

IADR (International Association of Dental Research) (2011) Study Shows Bone Fluoride Levels Not Associated with Osteosarcoma [Press Release]. http://www.iadr.org/files/public/JDRNewsRelease_Osteosarcoma.pdf [Accessed online June 1, 2012]

Jackson D, Weidmann S. Fluorine in human bone related to age and the water supply of different regions. *J Pathol Bacteriol*. 1958;76(2):451–459.

Katz, MH, 2011. *Multivariable Analysis - A Practical Guide for Clinicians and Public Health Researchers, 3rd edition*, Cambridge University Press.

Kim FM, Hayes C, Williams PL, et al. An assessment of bone fluoride and osteosarcoma. *J Dent Res*. 2011;90(10):1171–1176.

Li Y, Liang C, Slemenda CW, et al. Effect of long-term exposure to fluoride in drinking water on risks of bone fractures. *J Bone Miner Res*. 2001;16(5):932–9.

Mihashi M, Tsutsui T. Clastogenic activity of sodium fluoride to rat vertebral body-derived cells in culture. *Mutat Res*. 1996;368:7–13.

Mirabello L, Troisi RJ, Savage SA. International osteosarcoma incidence patterns in children and adolescents, middle ages and elderly persons. *Int J Cancer*. 2009;125(1):229–34.

Mirabello L, Pfeiffer R, Murphy G, et al. Height at diagnosis and birth-weight as risk factors for osteosarcoma. *Cancer Causes Control*. 2011;22(6):899–908.

National Research Council, Committee on Fluoride in Drinking Water (2006) Fluoride in Drinking Water: A Scientific Review of EPA's Standards. Washington DC: National Academies Press.

National Toxicology Program (1990) Toxicology and carcinogenesis studies of sodium fluoride (CAS No. 7681-49-4) in F344/N rats and B6C3F$_1$ mice (drinking water studies). Research Triangle Park NC: National Toxicology Program, National Institutes of Health, Report No. TR-393, NIH Publication No. 90–2848.

Savage SA, Mirabello L. Using epidemiology and genomics to understand osteosarcoma etiology. *Sarcoma*. 2011;2011:1 – 13.

Troisi R, Masters MN, Joshipura K, Douglass C, Cole BF, Hoover RN. Perinatal factors, growth and development, and osteosarcoma risk. *Br J Cancer*. 2006;95(11):1603–1607.

Rothman KJ. Induction and latent periods. *Am J Epidemiol*. 1981;114(2):253–259.

Rothman KJ, Greenland S. *Modern Epidemiology, 2nd Edition*. Lippincott Williams & Wilkins; 1998.

Zeiss EE, Hanley JA. Mantel-Haenszel techniques and logistic regression: always examine one's data first and don't overlook the simpler techniques. *Paediatr Perinat Epidemiol*. 1992;6(3):311–315.

Zipkin I, Mcclure FJ, Leone NC, Lee WA. Fluoride Deposition in Human Bones after Prolonged Ingestion of Fluoride in Drinking Water. *Public Health Rep*. 2015;73(8):732–740.