JOURNAL OF
EMERGING INVESTIGATORS

Article

Correlation between trihalomethane concentrations and various cancers in Massachusetts counties

Sashwat Ganeshwaran[1], Michael Romano[1]
[1]Acton-Boxborough Regional High School, Acton, Massachusetts

SUMMARY

Understanding the impact of water pollution on human health can guide new policies and technologies to improve human safety. Trihalomethanes (THMs) have become a prevalent pollutant in drinking water as the practice of chlorine disinfection has become common. As THMs are lipophilic (fat-soluble), they can accumulate in fatty tissues, including organs such as the liver and kidneys, raising concerns about long-term health effects. We examined the correlation between THM concentrations in drinking water and the rates of two cancer types: intrahepatic cholangiocarcinoma (ICC) and kidney and renal pelvis cancer. We hypothesized that a strong positive correlation would exist between total trihalomethane (TTHM) concentrations and the incidence and mortality rates for both cancers. To investigate this, we utilized data on annual mean TTHM concentrations in community water sources by Massachusetts county from the Centers for Disease Control (CDC) and age-adjusted cancer incidence and mortality rates from the Massachusetts Cancer Registry (MCR) for the years 2016-2020. Using this data, we found evidence of a positive correlation between TTHM concentrations and ICC incidence and mortality rates, as well as kidney and renal pelvis cancer incidence and mortality rates. Our results, combined with the results from other animal and human studies, raise concern that exposure to THMs, even at low levels, could contribute to an increase in incidence rates of intrahepatic bile and kidney and renal pelvis cancers.

INTRODUCTION

Trihalomethanes (THMs) are a class of disinfection byproducts formed when chlorine, a common disinfectant for drinking water, reacts with naturally occurring organic material (1). Chemically, THMs are created when three of the four hydrogen atoms of a methane molecule are replaced by halogen atoms like chlorine or bromine (1, 2). Around 70% of public water supplies in the United States are treated with chlorine, leaving residual THMs, which can enter the human body through three main routes: ingestion (drinking), inhalation (breathing), and dermal absorption (touching) (1). When showering or swimming in waters treated with chlorine, THMs can be absorbed through the skin, resulting in a THM concentration in the blood approximately two times greater than baseline (3). THMs have been associated with liver, kidney, and testicular diseases, along with an increased risk for colon, bladder, and rectal cancers (4). The most common THMs are chloroform, dibromochloromethane (DBCM), bromodichloromethane (BDCM), and bromoform; the sum of their concentrations is known as total trihalomethanes (TTHMs) (5). THMs, being lipophilic, tend to accumulate in fat; chloroform, for example, is found at high levels in fatty tissue and the liver following exposure (6).

The Centers for Disease Control and Prevention (CDC) describe kidney and renal pelvis cancer as a disease where abnormal cell growth occurs in the kidneys and their associated structures (7). The kidneys are a pair of organs in the abdomen that filter waste from the blood and play a critical role in regulating the homeostasis of chemicals in the body (8). The renal pelvis is a hollow, funnel-shaped cavity that collects urine from the kidneys before it passes into the ureter (9). In the United States, the 5-year relative survival rate for kidney and renal pelvis cancer is 77.6% (2014-2020) (10). In 2025, these cancers are projected to make up 4% of all new cancer incidences and 2.3% of all cancer-related deaths in the United States (11). Between 2007-2011, kidney and renal pelvis cancers accounted for 3.1% of all cancer incidence in Massachusetts (12). THMs have been associated with kidney damage, particularly in children and pregnant women (13). THMs can cause this damage by inhibiting the kidney's ability to filter blood and remove waste, leading to cumulative damage over time (14). In laboratory rats and mice, oral treatment of BDCM led to the development of kidney cancer (15).

Intrahepatic bile duct cancer, also known as intrahepatic cholangiocarcinoma (ICC), is a rare disease that occurs when malignant cells form in the small bile ducts within the liver (16, 17). These ducts collect bile, which helps break down fats, and transport it toward the gallbladder (18, 19). For individuals with ICC, the average 5-year survival rate was only 9% between 2012-2018 (20). In animal studies, exposure to THMs has been linked to the development of liver tumors (21).

Given the negative health impacts of THMs and their potential to accumulate in organs like the liver and kidneys, we explored the relationship between TTHM concentrations and the incidence and mortality of cancer rates in these areas. We also explored logarithmic models, as biological responses to toxins can exhibit non-linear relationships, such as plateauing effects at higher concentrations. We hypothesized that a positive correlation would exist between TTHM concentrations and the incidence and mortality rates for both kidney and renal pelvis cancer and ICC. The data showed a statistically significant positive correlation in both incidence and mortality for all the cancer types investigated. The results of this study highlight the importance of further

https://doi.org/10.59720/25-079

investigating strategies for reducing THM concentrations in drinking water, potentially including alternative disinfection methods.

RESULTS

To test our hypothesis that a positive correlation exists between TTHM concentrations and specific cancer rates, we conducted two main sets of analyses. First, we examined the correlation between TTHM concentrations and the incidence and mortality rates of kidney and renal pelvis cancer. Second, we performed a similar analysis for TTHM concentrations and the incidence and mortality rates of ICC. Both linear and logarithmic models were explored for these relationships.

When examining the relationship between TTHM concentrations and kidney and renal pelvis cancer in Massachusetts counties from 2016 to 2020, we found a moderate positive correlation for both incidence ($r = 0.422$, $p < 0.01$; **Figure 1A**) and mortality rates ($r = 0.415$, $p < 0.01$; **Figure 1B**). These results suggest that the annual mean concentration of TTHM has a statistically significant positive relationship with the age-adjusted rates of kidney and renal pelvis cancer.

Similarly, our analysis revealed a moderate positive correlation between TTHM concentrations and age-adjusted ICC incidence rates ($r = 0.468$, $p < 0.01$; **Figure 1C**) and mortality rates ($r = 0.395$, $p < 0.01$; **Figure 1D**). These statistically significant relationships provide strong evidence of a positive association between annual TTHM concentrations and the rates of ICC in Massachusetts.

While we initially focused on a linear regression model, we also found that a logarithmic model provided a better fit for three of the four outcomes. For kidney and renal pelvis cancer, the logarithmic model yielded correlation coefficients of 0.373 for incidence rates and 0.510 for mortality rates (**Figure 2A-B**). For ICC, the logarithmic model resulted in r-values of 0.525 for incidence and 0.555 for mortality rates (**Figure 2C-D**). All logarithmic models showed a statistically significant relationship ($p < 0.01$), and the higher r-values for kidney cancer mortality and both ICC metrics suggest that a non-linear model may better represent the dose-response relationship for these outcomes.

DISCUSSION

The evidence from all of our datasets supported our hypothesis that a positive correlation would exist between annual TTHM concentrations and the age-adjusted incidence and mortality rates of both kidney and renal pelvis cancer and ICC. This correlation might be explained by THMs accumulating in fatty tissue. Extrahepatic tissues, or tissues not part of the liver, receive higher relative doses of THMs from dermal absorption and inhalation compared to ingested THMs (22). As stated before, THMs absorbed through contact with water can result in THM concentrations up to two times the normal level (3).

The correlation observed between ICC rates and TTHM concentrations could be related to the physiology of the liver and biliary system. THMs are lipophilic (fat-soluble) compounds (6). Bile acids, synthesized in the liver and transported through the intrahepatic bile ducts, are crucial for emulsifying dietary fats (19, 23). It is plausible that THMs, accumulating in liver tissue or interacting with bile components due to their lipophilic nature, could lead to increased exposure

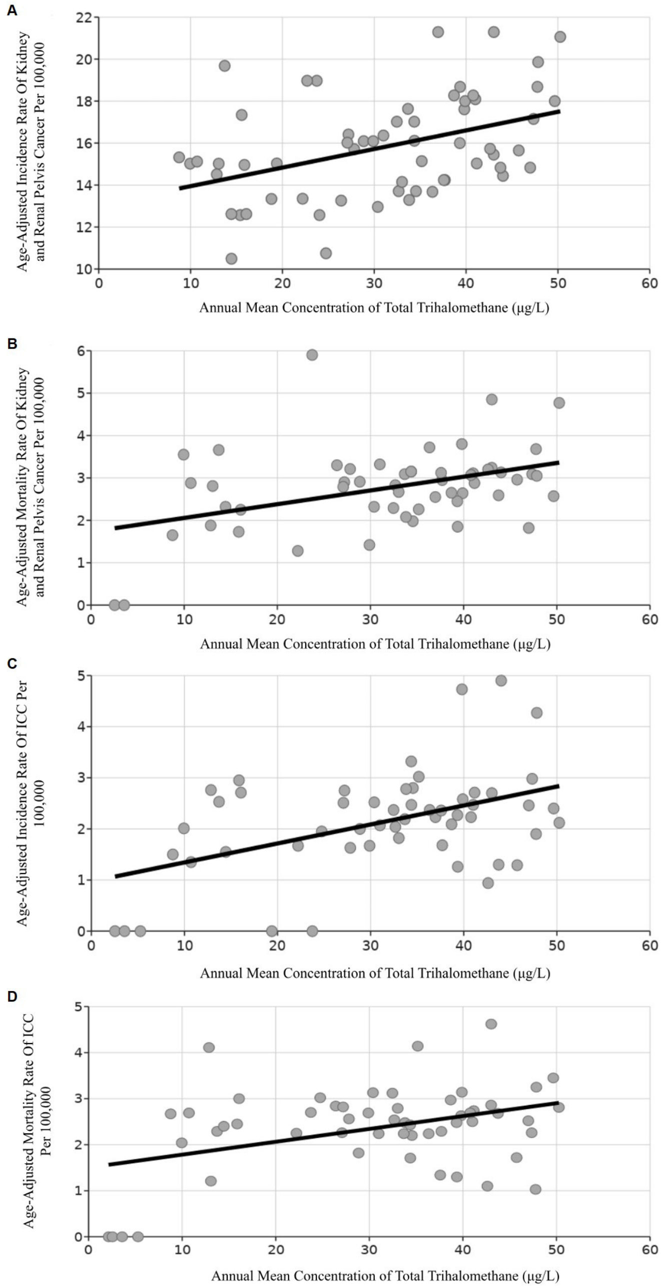

Figure 1: Positive linear correlation between annual mean total trihalomethane (TTHM) concentrations in community water sources and age-adjusted cancer incidence rates and mortality rates in Massachusetts counties (2016-2020). A) Each data point represents one county-year observation for kidney and renal pelvis cancer incidence rates ($n = 60$). The regression line is defined by the equation $y = 0.08855x + 13.06$ with a t-statistic of 3.5 and an $r^2 = 0.178$. **B)** Each data point represents one county-year observation for kidney and renal pelvis cancer mortality rates ($n = 52$). The regression line is defined by the equation $y = 0.03241x + 1.732$ with a t-statistic of 3.3 and an $r^2 = 0.172$. **C)** Each data point represents one county-year observation for intrahepatic cholangiocarcinoma (ICC)) incidence rates ($n = 53$). The regression line is defined by the equation $y = 0.03711x + 0.9733$ with a t-statistic of 3.8 and an $r^2 = 0.219$. **D)** Each data point represents one county-year observation for ICC mortality rates ($n = 56$). The regression line is defined by the equation $y = 0.02793x + 1.505$ with a t-statistic of 3.1 and an $r^2 = 0.156$.

https://doi.org/10.59720/25-079

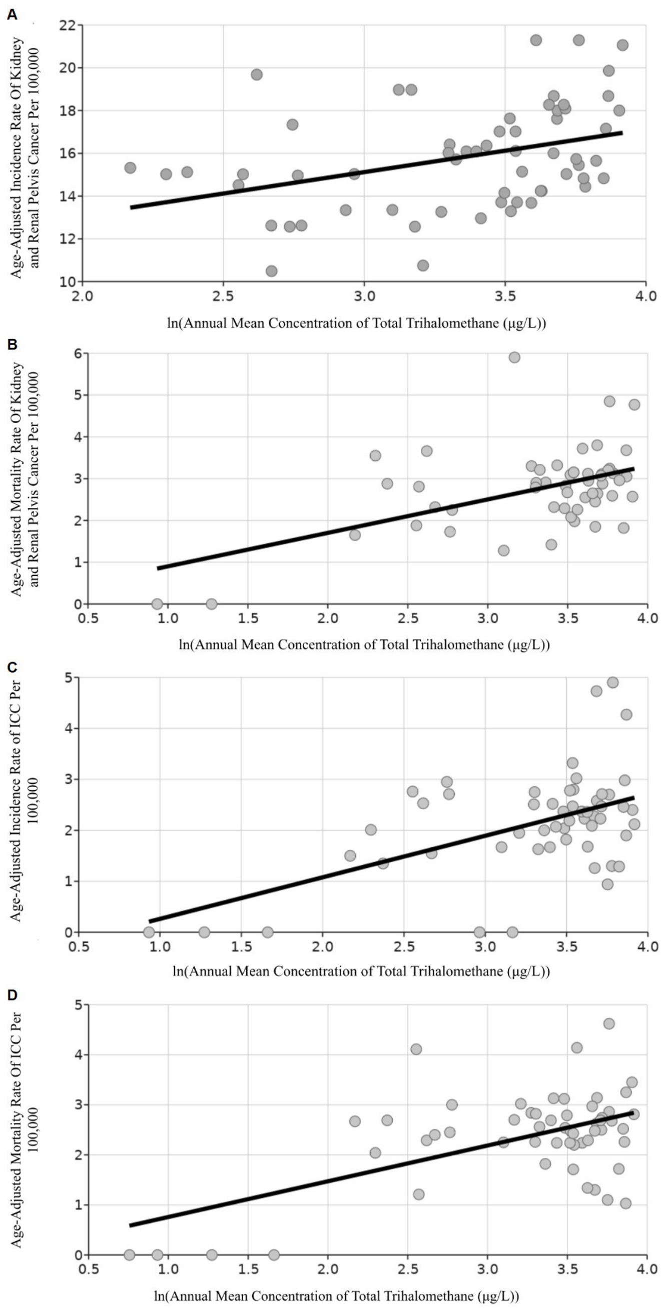

Figure 2: Positive logarithmic correlation between the natural log of annual mean total trihalomethane (TTHM) concentrations and age-adjusted cancer incidence rates and mortality rates in Massachusetts counties (2016-2020). A) Each data point represents one county-year observation for kidney and renal pelvis cancer incidence rates (n = 60). The regression line is defined by the equation y = 2.003x + 9.112 with a t-statistic of 3.1 and an r^2 = 0.139. **B)** Each data point represents one county-year observation for kidney and renal pelvis cancer mortality rates (n = 52). The regression line is defined by the equation y = 0.8004x + 0.1024 with a t-statistic of 4.3 and an r^2 = 0.260. **C)** Each data point represents one county-year observation for ICC incidence rates (n = 53). The regression is defined by the equation y = 0.815x – 0.551 with a t-statistic of 4.5 and an r2 = 0.276. **D)** Each data point represents one county-year observation for ICC mortality rates (n = 56). The regression line is defined by the equation y = 0.7139x + 0.04503 with a t-statistic of 4.453 and an r^2 = 0.308.

of the cells lining the intrahepatic bile ducts. The biliary tract's close anatomical and functional relationship with the liver, a primary site for metabolizing toxins like THMs, might further concentrate this exposure (22). This prolonged contact could contribute to the observed correlation.

The observation that logarithmic models in some cases provided an equal or better fit to the data (**Figures 5-8**) warrants further discussion. This pattern, where cancer rates appear to increase less steeply at higher TTHM concentrations, could be due to several biological phenomena. For instance, there might be a saturation point for the absorption or damaging effects of THMs within the target tissues. The body's capacity to process toxins like THMs is finite; at a certain point, the metabolic pathways responsible for detoxification can become saturated (25). Once such a saturation point is reached, further increases in external TTHM concentrations might not translate into proportionally higher internal damage, leading to the observed plateauing of cancer rates. Based on our data for ICC, this tapering appeared to begin at TTHM concentrations around 5-7.5 μg/L for incidence and 7.5-10 μg/L for mortality. Such non-linear dose-response relationships can occur if THMs have their greatest relative impact at low-to-medium exposure levels. If such saturation effects are at work, it would imply that for these cancers, public health interventions aimed at reducing TTHM exposure to levels below any potential inflection point could be impactful in mitigating risk.

The positive correlation between THM concentrations and kidney and renal pelvis cancer incidence and mortality rates could also be due to the kidneys' role in the body, filtering THMs from the blood and transporting urine. During the filtration process, kidney tissues could experience heightened exposure to THMs. Furthermore, the renal pelvis, a funnel-shaped structure that collects urine from the kidney and channels it to the ureter, would be in contact with urine containing concentrated THMs and their metabolites (9, 24). This continuous exposure of the urothelium lining the renal pelvis to these compounds could potentially lead to cellular damage, promote mutations over time, and contribute to cancer development (9). This chronic, prolonged exposure might explain why a logarithmic model better represents the data set compared to a linear model. It could suggest that even sustained exposure to low to medium levels of THMs might be sufficient to contribute to cancer risk over time. This could explain why very high levels of TTHM concentrations do not proportionally increase cancer incidence and mortality rates further.

Further research is needed to elucidate the biochemical pathways through which THMs might exert carcinogenic effects in both kidney and liver tissues. Understanding why these organs might be particularly susceptible compared to others, perhaps due to specific metabolic enzymes or transport mechanisms, also requires more investigation. In addition, this chronic, prolonged exposure might explain why a logarithmic model better represents the data set compared to a linear model. With prolonged exposure, TTHM concentrations do not have to be higher to cause cancer; instead, this could suggest that even sustained exposure to low to medium levels of THMs might be sufficient to contribute to cancer risk over time. This could explain why high levels of TTHM concentrations do not increase cancer incidence and mortality rates significantly.

https://doi.org/10.59720/25-079

While there is a positive correlation between annual TTHM concentration and the various cancers, THMs do not necessarily cause cancer. These cancers have multiple risk factors, and the observed associations could be influenced by various confounding variables not accounted for in this analysis. THMs may act as one of many contributing factors. This study is limited by its geographical scope and sample size; a larger, national study is needed to verify the results we observed. Because of this, it may be unrealistic to draw general conclusions beyond Massachusetts based on this study. While determining a definitive safe threshold for TTHM exposure concerning these cancers is beyond the scope of this study, our findings, especially the observed correlations even at the relatively low average TTHM concentrations found in some counties, raise questions about whether current regulatory limits are sufficiently protective for all over long exposure durations. To determine a precise concentration at which the relationship ceases to be significant, larger epidemiological studies, as well as animal and *in vitro* cell experimentation, would need to be performed to gather better results. Chlorine disinfection has provided microbially safe drinking water, dramatically reducing waterborne infectious diseases (26). The challenge lies in balancing these undeniable benefits against the potential long-term health risks from disinfection byproducts like THMs. This study contributes to the body of evidence suggesting that efforts to minimize disinfection byproduct formation, through optimizing chlorination processes or exploring alternative water treatment technologies, are warranted, provided that microbial safety is not compromised.

MATERIALS AND METHODS

Data Acquisition

To investigate the correlation between TTHM concentrations and cancer rates, we analyzed data from 2016 to 2020. To determine the average TTHM concentration for each Massachusetts county, we averaged the annual mean concentrations reported for all Community Water Systems (CWS) per person within that county. The data for all years used, 2016-2020, for TTHM concentrations were found using the National Environmental Public Health Tracking Network from the Centers for Disease Control and Prevention (27). From the Massachusetts Cancer Registry (MCR), using the MCR Web Query Tool, we obtained age-adjusted cancer mortality and incidence rates per 100,000 people from 2016-2020 per Massachusetts county (28). We used age-adjusted rates as they control for differences in age distribution across counties, a significant factor in cancer risk (29).

Data Availability and Exclusions

For TTHM concentrations, data for Nantucket County (2016-2020) were not available as no Community Water Sources (CWS) were reported for this county in the CDC database during the study period. For cancer data, kidney and renal pelvis cancer incidence rates for Dukes County (2016-2020) were unavailable due to confidentiality concerns, given the small case numbers. Kidney and renal pelvis cancer mortality rates were unavailable for Dukes County (2016, 2018, 2020), Franklin County (2016-2020), and Hampshire County (2018-2020) due to confidentiality concerns. For ICC incidence rates, data were unavailable for Barnstable County (2017), Berkshire County (2017), Dukes County (2016, 2020), Franklin County (2016-2019), and Hampshire County (2016, 2018-2020) due to confidentiality concerns. For ICC mortality rates, data were unavailable for Berkshire County (2016, 2018), Dukes County (2020), Franklin County (2016-2019), and Hampshire County (2018-2019) due to confidentiality concerns.

Statistical Analysis

In order to determine if a correlation existed between TTHM concentrations and the various cancers, we performed a linear regression analysis to calculate Pearson Correlation Coefficient R-values and p-values. To explore a potential logarithmic relationship between THMs and the various cancers, we transformed the values of the independent variable, annual mean THM concentrations, using the natural logarithm (ln). A linear regression test was then run to determine r-values. Data Classroom was used to perform statistical analyses and generate figures (30).

Received: April 4, 2025
Accepted: June 10, 2025
Published: September 8, 2025

REFERENCES (30)

“Trihalomethanes (Chlorine Disinfection Byproducts).” Delaware Health and Social Services - State of Delaware. Accessed 15 Nov. 2024. https://dhss.delaware.gov/dhss/dph/files/Trihalomethanes_FAQ_PUB_ENG_0323.pdf.

Hood, Ernie. “Tap Water and Trihalomethanes: Flow of Concerns Continues.” Environmental Health Perspectives, vol. 113, no. 7, July 2005, p. A474, https://doi.org/10.1289/ehp.113-a474.

Nuckols, John R., et al. “Influence of Tap Water Quality and Household Water Use Activities on Indoor Air and Internal Dose Levels of Trihalomethanes.” Environmental Health Perspectives, vol. 113, no. 7, July 2005, pp. 863–870, https://doi.org/10.1289/ehp.7141.

“Trihalomethanes Fact Sheet 2017.” Florida Department of Health, Bureau of Environmental Health. Accessed 19 Nov. 2024. https://www.floridahealth.gov/environmental-health/drinking-water/_documents/thms-fs-2017.pdf.

“Trihalomethanes - HSE.ie.” Health Services Executive. Accessed 18 Nov. 2024. https://www.hse.ie/eng/health/hl/water/drinkingwater/trihalomethanes/.

Gad, Shayne C. “Trihalomethanes.” Encyclopedia of Toxicology 4th Edition, 4th ed., vol. 9, Elsevier Inc., 2023, pp. 615–619.

“Kidney Cancer Basics.” Center for Disease Control and Prevention. Accessed 29 Nov. 2024. https://www.cdc.gov/kidney-cancer/about/index.html.

“NCI Dictionary of Cancer Terms.” National Cancer Institute. Accessed 1 Dec. 2024. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/kidney.

“NCI Dictionary of Cancer Terms.” National Cancer Institute. Accessed 2 Dec. 2024. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/renal-pelvis.

“Cancer Stat Facts: Kidney and Renal Pelvis Cancer.” National Cancer Institute, Surveillance, Epidemiology, and End Results Program. Accessed 18 Aug. 2024. https://seer.cancer.gov/statfacts/html/kidrp.html.

Siegel, Rebecca L., et al. “Cancer Statistics, 2025.” CA: A Cancer Journal for Clinicians, vol. 75, no.1, Jan. 2025, pp.

https://doi.org/10.59720/25-079

5027. https://doi.org/10.3322/caac.21871.
"Risk Factor Information for Kidney and Renal Pelvis Cancers." Massachusetts Department of Public Health. Accessed 7 Dec. 2024. https://matracking.ehs.state.ma.us/Health-Data/Cancer/factsheets/Kidney_and_Renal_Pelvis.pdf.
"Total Trihalomethane in Drinking Water: Information for Consumers." Massachusetts Department of Environmental Protection. Accessed 7 Dec. 2024. https://www.mass.gov/doc/factsheet-total-trihalomethane-in-drinking-water-information-for-consumers/download.
Kroll, R.B, et al. "Characterization of Trihalomethane (THM)-Induced Renal Dysfunction in the Rat. I: Effects of THM on Glomerular Filtration and Renal Concentrating Ability." Archives of Environmental Contamination and Toxicology, vol. 27, no. 1, 1 July 1994, pp. 1-6, https://doi.org/10.1007/bf00203878.
"Bromodichloromethane: CAS No. 75-27-4." 15th Report on Carcinogens. Research Triangle Park (NC): National Toxicology Program; 21 Dec. 2021. Accessed 25 Aug. 2025. Bromodichloromethane. https://doi.org/10.22427/NTP-OTHER-1003
"NCI Dictionary of Cancer Terms." National Cancer Institute. Accessed 17 Dec. 2024. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/intrahepatic-bile-duct.
"What Is Bile Duct Cancer (Cholangiocarcinoma)?" National Cancer Institute. Accessed 17 Dec. 2024. https://www.cancer.gov/types/liver/bile-duct-cancer.
"What Are Bile Ducts?" Cleveland Clinic. Accessed 21 Dec. 2024. https://my.clevelandclinic.org/health/body/24523-bile-duct.
Hundt, M. A., & Basit, H. "Physiology, Bile." In: StatPearls. StatPearls Publishing, 2024. https://www.ncbi.nlm.nih.gov/books/NBK470209/.
"Bile Duct Cancer Survival Rates | Cholangiocarcinoma Survival Rates." American Cancer Society. Accessed 17 Dec. 2024. https://www.cancer.org/cancer/types/bile-duct-cancer/detection-diagnosis-staging/survival-by-stage.html.
"Trihalomethanes (THMs) - the Center for Health, Environment & Justice." The Center for Health, Environment & Justice. Accessed 21 Dec. 2024. https://chej.org/trihalomethanes-thms.
"Trihalomethanes in Drinking-water." World Health Organization. Accessed 21 Dec. 2024. https://www.who.int/docs/default-source/wash-documents/wash-chemicals/trihalomethanes.pdf.
"Biliary System Anatomy and Functions." Johns Hopkins Medicine. Accessed 24 Dec. 2024. https://www.hopkinsmedicine.org/health/conditions-and-diseases/biliary-system-anatomy-and-functions.
"Kidneys | SEER Training." Epidemiology and End Results Program Training Modules, National Cancer Institute. Accessed 26 Dec. 2024. https://training.seer.cancer.gov/anatomy/urinary/components/kidney.html.
Gillette, J. R. "A Perspective on the Role of Chemically Reactive Metabolites of Foreign Compounds in Toxicity: I. Correlation of Changes in Covalent Binding of Reactive Metabolites with Changes in the Incidence and Severity of Toxicity." Biochemical Pharmacology, vol. 23, no. 20, 15 Oct. 1974, pp. 2785–2794, https://doi.org/10.1016/0006-2952(74)90052-5.
"The Disinfection of Drinking Water." Drinking Water and Health: Volume 2. National Research Council (US) Safe Drinking Water Committee. National Academies Press (US); 1980. Accessed 25 Aug. 2025. https://www.ncbi.nlm.nih.gov/books/NBK234590/.
"National Environmental Public Health Tracking Network Data Explorer." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention. Accessed 2 Nov. 2024. https://ephtracking.cdc.gov/DataExplorer/.
"Massachusetts Cancer Registry Web Query Tool." Massachusetts Cancer Registry. Accessed 2 Nov. 2024. https://www.cancer-rates.com/ma/.
"Age-Adjusted Rates." New York State Department of Health. Accessed 25 Aug. 2025. https://www.health.ny.gov/diseases/chronic/ageadj.htm.
Dataclassroom. Accessed 14 Nov. 2024. https://dataclassroom.com/en.