Journal of Research High School

Vol. 2026 (7) 570 – 577
ISSN 2688-3651 [online]

JRHS Outstanding Research Paper Award

Drinking-Water PFAS Concentrations in New Jersey: Sociodemographic Analysis and Structure-Related Contaminant Correlations

Albert Lee [1 *]

[1]Horace Mann School, Bronx, NY, USA
*Corresponding Author: alberthlee27@gmail.com

Advisor: Dr. Sang Park, spark1@ggc.edu

Received February 22, 2026; Revised April 7, 2026; Accepted April 20, 2026

Abstract

Perfluoroalkyl and polyfluoroalkyl substances (PFAS) are persistent environmental contaminants of increasing public health concern. Previous national studies have reported associations between PFAS exposure and sociodemographic variables. This study investigated relationships between concentrations of three PFAS pollutants—PFOA, PFOS, and PFBS—and selected sociodemographic indicators, including population served, education level, and median household income, using publicly available drinking-water data from Community Water Systems in New Jersey. Data were separated by contaminant, and outliers were removed using the Grubbs test ($\alpha = 0.05$). Pearson correlation and linear regression analyses were conducted to calculate Pearson correlation coefficients, R^2 values, adjusted R^2 values, and associated p-values. Results showed no consistent or statistically significant associations between PFAS concentrations and sociodemographic variables, except for a weak but statistically significant positive correlation between household income and PFOA concentration; however, the low R^2 value (0.0997) indicated limited explanatory power. In contrast, stronger correlations were observed among certain PFAS compounds, and the strength of these inter-compound relationships corresponded with similarities in carbon chain length and functional groups. These findings suggest that, within this statewide dataset, PFAS concentration patterns may be more strongly associated with compound-specific chemical characteristics than with sociodemographic factors.

Keywords: PFAS, Sociodemographic factors, Pearson correlation, Drinking water contamination

1. Introduction

Perfluoroalkyl and polyfluoroalkyl substances (PFAS) are a large class of synthetic chemicals that have been widely used since the mid-20th century in industrial processes and consumer products. Their unique molecular structure, consisting of carbon chains fully bonded to fluorine atoms, creates some of the strongest bonds in organic chemistry, giving PFAS exceptional thermal stability, resistance to degradation, and the ability to repel water and oil (Buck et al., 2011). These same properties, however, make PFAS extremely persistent in the environment, "earning them the name 'forever chemicals' (Wang et al., 2017)."

PFAS production began expanding rapidly in the late 1940s and early 1950s as companies such as 3M and DuPont introduced chemicals, including perfluorooctanoic acid (PFOA), polytetrafluoroethylene (Teflon), and perfluorooctane sulfonic acid (PFOS), one of the earliest and most commercially important PFAS chemicals (American Chemical Society, n.d.). By the 1960s, PFAS had become common in cookware coatings, stain-repellent textiles, firefighting foams, and numerous industrial applications. During this period, few environmental rules existed, and PFAS waste was often released into waterways or buried without attention to its long-term effects, allowing these chemicals to accumulate in soil, groundwater, and surface water, many of which serve as drinking-water sources.

Although internal company documents indicated awareness of PFAS toxicity as early as the 1970s, the health risks were not widely known to the public until legal cases in the late 1990s and early 2000s revealed extensive

Vol. 2026 (7) 570 – 577
ISSN 2688-3651 [online]

contamination and harmful effects (Lerner, 2021). Later studies have found that exposure to some PFAS can raise the risk of cancer, slow development in children, weaken the immune system, and cause reproductive problems. In response, both federal and state agencies have implemented limits on major PFAS compounds (Grandjean & Clapp, 2015; Barry et al., 2013). In New Jersey, current enforceable standards include 14 parts per trillion (ppt) for PFOA and 13 ppt for PFOS, although Perfluorobutane sulfonic acid (PFBS) and many other PFAS remain unregulated (New Jersey DEP, 2020).

PFAS contamination is now recognized as a widespread national issue. A 2023 study by the U.S. Geological Survey estimated that at least 45 percent of tap water in the United States contains one or more PFAS chemicals (US Geological Survey, 2023). However, exposure is not evenly distributed. Recent studies show that certain communities, particularly Black and Hispanic or Latino populations, are more likely to live near PFAS sources or have detectable PFAS in drinking water. Factors such as income level, population density, industrial history, and education may contribute to these disparities (Liddie et al., 2023; Hu et al., 2016). These variables serve as proxies for socioeconomic status, community resources, and potential exposure vulnerability.

Despite growing national research on PFAS exposure and environmental justice, no comprehensive statewide analysis has examined how PFAS concentrations vary with social and demographic factors in New Jersey. Addressing this gap is important because the state has a long industrial history, diverse communities, and varying levels of regulatory action.

To address this gap, the present study analyzes publicly available data on PFOA, PFOS, and PFBS concentrations across New Jersey and investigates their relationships with demographic and socioeconomic variables, including income, education, and population characteristics. Understanding these patterns may help identify communities at elevated risk and support more equitable environmental policy decisions.

This study tested the hypothesis that PFAS concentrations are significantly associated with sociodemographic indicators at the Community Water System level in New Jersey. Additionally, relationships among PFAS compounds were evaluated to assess potential structure-related co-occurrence patterns.

2. Methods

2.1 Data Collection and Process

PFAS testing data from 491 Community Water Systems (CWSs) were obtained from the New Jersey Department of Environmental Protection (NJDEP) for the years 2022–2024. A master spreadsheet was created to organize the dataset, including fields for water utility name, service location, estimated population served, and measured PFAS concentrations (ppt). Data were manually transferred into the spreadsheet to verify completeness and ensure accurate alignment across variables.

The dataset was separated into three subsets corresponding to the individual contaminants, PFOA, PFOS, and PFBS. CWSs lacking measurements for a specific contaminant were excluded from that subset. Outliers were identified and removed using the Grubbs test ($\alpha = 0.05$) using Equation 1 (Harris, 2009). Outliers identified by the Grubbs test were limited in number and were removed cautiously to minimize undue influence on statistical results.

$$G = \frac{|questionable\ value - \bar{x}|}{s} \tag{1}$$

where $\bar{x}$ and s indicate the data set's average and the standard deviation, respectively.

Demographic information was obtained from the U.S. Census Bureau's American Community Survey (ACS 2016–2020). ACS indicators, including high school graduation rate, bachelor's degree attainment, and median household income, were spatially joined to CWS service area boundaries to estimate the sociodemographic characteristics of populations served. These variables were matched to their corresponding CWS entries for analysis.

Vol. 2026 (7) 570 – 577
ISSN 2688-3651 [online]

2.2 Statistical Analysis

Comparisons were conducted between each PFAS contaminant (PFOA, PFOS, PFBS) and the sociodemographic variables (population served, high school graduation rate, bachelor's degree attainment, and median household income) using the Pearson correlation coefficient (ρ), as expressed in Equation 2.

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \tag{2}$$

Where x_i and y_i are individual paired observations, $\bar{x}$ and $\bar{y}$are the sample means, s_x and s_y represent the standard deviations of the *x* and *y* variables, respectively, and *n* is the sample size.

Statistical significance for all tests was evaluated at an alpha level of 0.05. The null hypothesis for each comparison stated that no linear correlation exists between the PFAS concentration and the sociodemographic factor of interest. This corresponds to $H_0{:}r = 0$, while the alternative hypothesis is $H_A{:}r \neq 0$. The Pearson correlation and t-test analyses assume (1) approximate normality of the data, (2) independence of observations, and (3) linear relationships between variables for correlation analysis.

Correlation coefficients and associated *p*-values were obtained using the linear regression module within the Microsoft Excel Data Analysis Toolpak, which computes *r*, the regression slope, and the corresponding significance test statistics.

Additional analyses were conducted to assess whether statistically significant differences existed between the mean concentrations of the three PFAS contaminants (PFOA, PFOS, and PFBS). Pairwise comparisons were performed using two-sample Student's *t*-tests with pooled variance at the 95% confidence level. For each contaminant pair, the *t*-statistic was calculated using the pooled standard deviation formula shown in Equation 3 (Harris, 2009).

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \quad \text{where} \quad s_{pooled} = \sqrt{\frac{s_1^2(n_1-1) + s_2^2(n_2-1)}{n_1 + n_2 - 2}} \tag{3}$$

Where denotes the pooled standard deviation, s_1 and s_2 are standard deviations of the two samples, and n_1 and n_2 are their respective sample sizes.

The *t*-statistic for each pair was then compared with the corresponding critical *t*-value to determine whether the difference in mean concentrations was statistically significant. When the *t*-statistic exceeded the critical *t*-value, the null hypothesis of no difference in mean concentrations was rejected.

It should be noted that the sample sizes varied across analyses due to differences in data availability for each PFAS compound and sociodemographic variable. Community Water Systems lacking measurements for a given contaminant or missing demographic data were excluded from that specific analysis. This variation in sample size may affect statistical power and comparability across models.

3. Results

3.1 PFAS Concentrations and Sociodemographic Variables

Pearson correlation analyses were conducted to evaluate relationships between PFAS concentrations (PFOA, PFOS, PFBS) and sociodemographic indicators, including population served, percentage of high school graduates, percentage of bachelor's degree attainment, and median household income.

Table 1 summarizes the regression statistics for PFOA using each sociodemographic variable as a single predictor. As shown in Table 1, household income demonstrated a statistically significant positive correlation with PFOA concentration ($r = 0.3157$, $p = 0.0034$). The R^2 value was 0.0997, indicating that approximately 10% of the variation in PFOA concentration is explained by income. No statistically significant correlations were observed for population served, high school graduation rate, or bachelor's degree attainment ($p > 0.05$).

Vol. 2026 (7) 570 – 577
ISSN 2688-3651 [online]

Table 1. Various statistical coefficients for PFOA using a single predictor regression model.

	Population	% high school graduates	% college graduates	Household Income
Number of Sample	84	74	74	84
Pearson Correlation Coefficient	0.1684	0.0965	0.1730	0.3157
R^2	0.0284	0.0093	0.0299	0.0997
Adjusted R^2	0.0165	-0.0044	0.0165	0.0887
Standard Error	3.6398	3.8155	3.7755	3.5037
P-value	0.1256	0.4133	0.1404	0.0034
Average PFOA Amount (ppt)	6.71 ± 3.18			

Table 2 presents the regression results for PFOS. As shown in Table 2, none of the sociodemographic variables exhibited statistically significant correlations with PFOS concentration (all $p > 0.05$). Correlation coefficients were low, ranging from 0.0213 to 0.1845, and R^2 values were below 0.05.

Table 2. Various statistical coefficients for PFOS using a single predictor regression model.

	Population	% high school graduates	% college graduates	Household Income
Number of Sample	50	43	43	50
Pearson Correlation Coefficient	0.1845	0.0516	0.0323	0.0213
R^2	0.0340	0.0027	0.0010	0.0005
Adjusted R^2	0.0139	-0.0217	-0.0233	-0.0204
Standard Error	2.4371	2.1065	2.1082	2.4791
P-value	0.2000	0.7423	0.8373	0.8832
Average PFOS amount (ppt)	6.18 ± 1.49			

Table 3 displays the regression statistics for PFBS. As shown in Table 3, no statistically significant relationships were identified between PFBS concentration and any sociodemographic variable (all $p > 0.05$). R^2 values were below 0.02, and several adjusted R^2 values were negative.

Table 3. Various statistical coefficients for PFBS using a single predictor regression model.

	Population	% high school graduates	% college graduates	Household Income
Number of Sample	32	30	30	32
Pearson Correlation Coefficient	0.1043	0.0153	0.0848	0.0089
R^2	0.0109	0.0002	0.0072	0.0001
Adjusted R^2	-0.0221	-0.0355	-0.0283	-0.0333
Standard Error	2.0395	2.0416	2.0345	2.0507
P-value	0.5698	0.9361	0.6558	0.9613
Average PFBS amount (ppt)	4.87 ± 2.01			

Figure 1 illustrates the relationship between household income and PFOA concentration. The scatterplot shows a slight upward trend consistent with the positive Pearson correlation coefficient reported in Table 1. However, the dispersion of data points indicates substantial variability around the regression line.

3.2 Relationships Among PFAS Compounds

Relationships among PFOA, PFOS, and PFBS concentrations were evaluated using linear regression analysis. Table 4 summarizes the correlation coefficients, R^2 values, adjusted R^2 values, and p-values for each pairwise comparison among the three PFAS compounds. As shown in Table 4, a statistically significant moderate positive correlation was observed between PFOA and PFOS ($r = 0.6444$, $R^2 = 0.4152$, $p = 6.78 \times 10^{-4}$). A strong positive

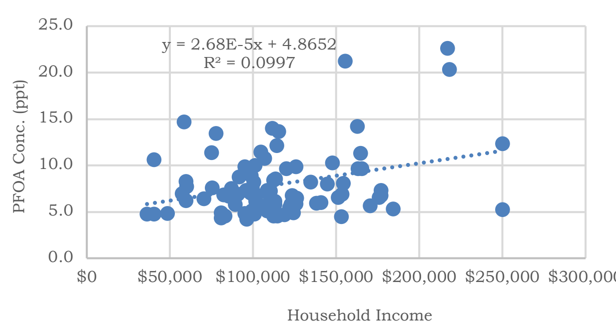

Figure 1. The relationship between PFOA concentration and household income.

correlation was observed between PFOS and PFBS ($r = 0.8721$, $R^2 = 0.7605$, $p = 0.0010$). In contrast, no statistically significant relationship was found between PFOA and PFBS ($r = 0.0036$, $R^2 = 1.31 \times 10^{-5}$, $p = 0.9911$).

Table 4. Statistical coefficients for the relationships among three PFAS contaminants.

	PFOA vs. PFOS	PFOA vs. PFBS	PFOS vs. PFBS
Number of Sample	24	12	10
Pearson Correlation Coefficient	0.6444	0.0036	0.8721
R^2	0.4152	1.31×10^{-5}	0.7605
Adjusted R^2	0.3886	-0.1000	0.7306
Standard Error	1.6091	2.1826	0.9358
P-value	6.78×10^{-4}	0.9911	0.0010

Figure 2 shows the scatterplot and regression line for PFOA versus PFOS concentrations. The data points display a positive linear trend consistent with the R^2 value reported in Table 4. Figure 3 shows the scatterplot and regression line for PFOS versus PFBS concentrations. The clustering of data points along the regression line reflects the strong positive relationship summarized in Table 4.

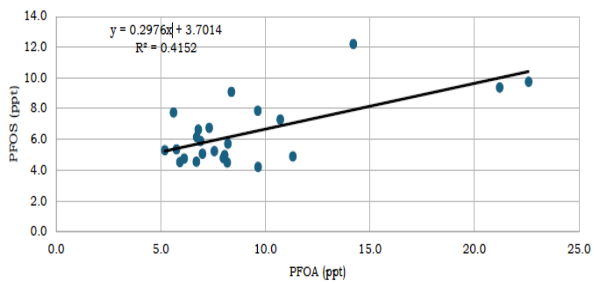

Figure 2. Scatterplot showing linear regression relationship between PFOA and PFOS concentrations in New Jersey Community Water Systems.

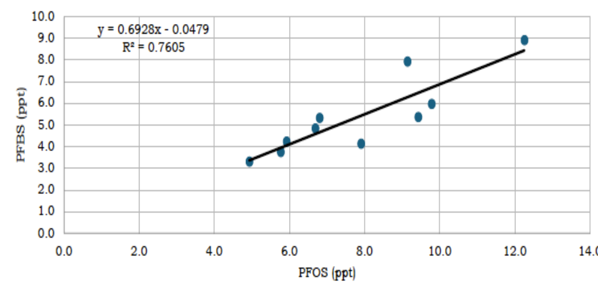

Figure 3. Scatterplot showing linear regression relationship between PFOS and PFBS concentrations in New Jersey Community Water Systems.

Figure 4 shows the scatterplot for PFOA versus PFBS. The data points are widely dispersed with no visible linear pattern, consistent with the near-zero correlation coefficient and non-significant p-value.

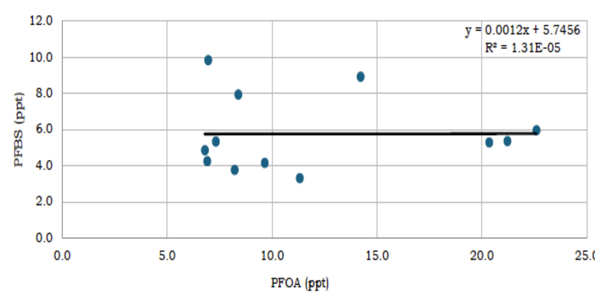

Figure 4. Scatterplot showing linear regression relationship between PFOA and PFBS concentrations in New Jersey Community Water Systems.

3.3 Differences in Mean Concentrations

Pairwise two-sample Student's t-tests with pooled variance were conducted for towns with at least three measurements per contaminant. The results indicated that statistically significant differences in mean concentrations were observed in 29.2% of towns for PFOA vs. PFOS, 75.0% for PFOA vs. PFBS, and 40.0% for PFOS vs. PFBS.

4. Discussion

4.1 Sociodemographic Factors and PFAS Distribution

The results indicate that sociodemographic indicators were not strong predictors of PFAS concentrations in this statewide dataset. Although household income showed a statistically significant positive relationship with PFOA concentration, the low R^2 value suggests limited explanatory strength. The absence of consistent associations across PFOS and PFBS further supports the conclusion that socioeconomic variables alone do not explain PFAS concentration patterns in New Jersey Community Water Systems. The absence of strong socioeconomic correlations may reflect differences in industrial history, regulatory enforcement, groundwater hydrology, or spatial distribution of contamination sources within the state. These findings suggest that PFAS exposure patterns in New Jersey may not follow the same sociodemographic trends reported in broader national datasets. Additionally, the use of aggregated data at the Community Water System level and variability in sample sizes may limit the ability to detect localized or subtle associations.

4.2 Structure-Based Correlation Patterns

In contrast to the weak sociodemographic associations, stronger statistical relationships were observed among certain PFAS compounds. As summarized in Table 4, moderate correlation was found between PFOA and PFOS ($R^2 = 0.4152$), and a strong correlation was observed between PFOS and PFBS ($R^2 = 0.7605$). These relationships are visually represented in Figures 2 and 3, respectively. Figure 2 demonstrates a positive linear trend between PFOA and PFOS concentrations, whereas Figure 3 shows a more tightly clustered distribution of data points along the regression line, consistent with the higher R^2 value. Conversely, Figure 4 shows the absence of a visible linear pattern between PFOA and PFBS, consistent with the near-zero correlation coefficient reported in Table 4. The observed correlations among PFAS compounds may reflect both shared contamination sources and similarities in chemical structure. For example, PFOS and PFBS are often associated with similar industrial applications, which may lead to co-release into the environment. At the same time, their shared functional groups may influence environmental behavior, including mobility and persistence. Therefore, the observed correlations likely arise from a combination of common sources and structural similarities. Structural similarity among PFAS compounds is environmentally significant because it influences key properties such as hydrophobicity, sorption behavior, and mobility in aquatic systems. Compounds with similar functional groups or chain lengths may exhibit comparable partitioning and persistence, increasing the likelihood of co-occurrence in water systems. To interpret these patterns, the molecular structures of the three PFAS compounds are presented in Table 5.

Table 5. Chemical structures of three PFAS pollutants studied in this paper.

PFOA	$CF_3 - CF_2 - CF_2 - CF_2 - CF_2 - CF_2 - CF_2 – CO_2H$
PFOS	$CF_3 - CF_2 - CF_2 - CF_2 - CF_2 - CF_2 - CF_2 - CF_2 – SO_3H$
PFBS	$CF_3 - CF_2 - CF_2 - CF_2 – SO_3H$

PFOA and PFOS share an eight-carbon fully fluorinated backbone, which may contribute to similar environmental persistence and co-occurrence patterns. PFOS and PFBS share a sulfonic acid functional group, which may influence polarity and mobility in aquatic systems. The absence of correlation between PFOA and PFBS corresponds to differences in both carbon chain length and functional group. These structural distinctions may contribute to differences in transport, solubility, and environmental partitioning.

Although statistical correlation does not establish causation, the alignment between structural similarity and correlation strength suggests that chemical properties may influence PFAS co-occurrence patterns within drinking-water systems.

4.3 Environmental and Monitoring Implications

The results indicate that PFAS concentration patterns in this statewide dataset appear to be more strongly associated with inter-compound relationships than with sociodemographic variables. The strong correlation between PFOS and PFBS (Figure 3) and the moderate trend between PFOA and PFOS (Figure 2) suggest that detection of one compound may signal the presence of structurally related compounds within a water system.

From a monitoring perspective, understanding structural relationships among PFAS compounds may improve assessment of co-occurrence risks and inform targeted analytical strategies. These findings also emphasize the importance of region-specific investigations, as statewide patterns may differ from national-level trends reported in previous research.

4.4 Study Limitations

Several limitations should be acknowledged. First, the analysis relied on publicly available data aggregated at the Community Water System level, which may obscure finer-scale geographic variability. Second, the study evaluated only three PFAS compounds, whereas many additional PFAS species may be present. Third, sample sizes varied across analyses and were relatively small, particularly for inter-compound comparisons in Table 4 (n = 10–24, with some comparisons as low as n = 10–12), which may reduce statistical power and increase sensitivity to variability. As a result, the reported correlations should be interpreted with caution. Fourth, a temporal mismatch exists between sociodemographic data (ACS 2016–2020) and PFAS measurements (2022–2024), which may introduce uncertainty. Finally, statistical analyses assume normality and independence, and outliers identified using the Grubbs test were limited in number and removed cautiously to minimize undue influence on statistical results. However, outliers in environmental datasets may represent true contamination hotspots rather than measurement error, and their exclusion may reduce variability and potentially underestimate localized contamination.

5. Conclusion

This study examined the relationships between three PFAS contaminants (PFOA, PFOS, and PFBS) and selected sociodemographic indicators across Community Water Systems in New Jersey. Statistical analyses showed that socioeconomic variables, including population served, education level, and median household income, were not consistent or strong predictors of PFAS concentrations. Although a statistically significant positive association was observed between household income and PFOA concentration, the low R^2 value indicates limited explanatory strength.

Unlike certain nationwide studies reporting socioeconomic disparities in PFAS exposure (Liddie et al., 2023; Hu et al., 2016), the present statewide analysis did not identify consistent or strong associations between PFAS concentrations and sociodemographic indicators in New Jersey.

In contrast, stronger relationships were observed among certain PFAS compounds themselves. Moderate correlation was found between PFOA and PFOS, and a strong correlation was observed between PFOS and PFBS. The absence of correlation between PFOA and PFBS further highlights differences in distribution patterns among these compounds. These inter-compound relationships appear to align with similarities in carbon chain length and functional groups, suggesting that chemical structure may influence PFAS co-occurrence in drinking-water systems.

The findings indicate that, within this statewide dataset, PFAS concentration patterns may be more strongly associated with compound-specific chemical characteristics than with sociodemographic variables. From a monitoring and regulatory perspective, understanding structural relationships among PFAS compounds may improve assessment of co-occurrence risks and inform targeted testing strategies. Future research incorporating hydrological factors, contamination source mapping, and multivariable modeling may further clarify the environmental processes influencing PFAS distribution across drinking-water systems.

References

Barry, V., Winquist, A., & Steenland, K. (2013). Perfluorooctanoic acid (PFOA) exposures and incident cancers among adults living near a chemical plant. *Environmental Health Perspectives*, 121(11-12), 1313-1318. https://pubmed.ncbi.nlm.nih.gov/24007715/

Brendel, S., et al. (2018). Short-chain perfluoroalkyl acids: Environmental concerns and a regulatory strategy under REACH. *Environmental Sciences Europe*, 30(9). https://doi.org/10.1186/s12302-018-0134-4

Buck, R. C., et al. (2011). Perfluoroalkyl and polyfluoroalkyl substances in the environment: Terminology, classification, and origins. *Integrated Environmental Assessment and Management*, 7(4), 513-541. https://doi.org/10.1002/ieam.258

Communications and Publishing. (2023, July 3). *Tap water study detects PFAS 'forever chemicals' across the US*. U.S. Geological Survey. Retrieved January 18, 2026, from https://www.usgs.gov/news/national-news-release/tap-water-study-detects-pfas-forever-chemicals-across-us

EPA PFAS rule. (n.d.). New Jersey Department of Environmental Protection. Retrieved January 18, 2026, from https://dep.nj.gov/pfas/epa-pfas-rule/

Vol. 2026 (7) 570 – 577
ISSN 2688-3651 [online]

Giesy, J. P., & Kannan, K. (2002). Peer reviewed: Perfluorochemical surfactants in the environment. *Environmental Science & Technology*, 36(7), 146A-152A. https://doi.org/10.1021/es022253t

Grandjean, P., & Clapp, R. (2014). Changing interpretation of human health risks from perfluorinated compounds. *Public Health Reports*, 129(6), 482-485. https://doi.org/10.1177/003335491412900605

Harris, D. C. (2009). *Exploring chemical analysis* (4th ed.). W. H. Freeman and Company.

Hu, X. C., et al. (2016). Detection of poly- and perfluoroalkyl substances (PFASs) in U.S. drinking water linked to industrial sites, military fire training areas, and wastewater treatment plants. *Environmental Science & Technology Letters*, 3(10), 344-350. https://doi.org/10.1021/acs.estlett.6b00260

Lerner, S. (2024, May 20). *Toxic gaslighting: How 3M executives convinced a scientist the forever chemicals she found in human blood were safe*. Propublica. Retrieved January 18, 2026, from https://www.propublica.org/article/3m-forever-chemicals-pfas-pfos-inside-story

Liddie, J. M., Schaider, L. A., & Sunderland, E. M. (2023). Sociodemographic factors are associated with the abundance of PFAS sources and detection in U.S. community water systems. E*nvironmental Science & Technology*, 57(21), 7902-7912. https://doi.org/10.1021/acs.est.2c07255

Perfluorooctanesulfonic acid. (2019, January 28). American Chemical Society. Retrieved January 18, 2026, from https://www.acs.org/molecule-of-the-week/archive/p/perfluorooctanesulfonic-acid.html

Wang, Z., et al. (2017). A never-ending story of per- and polyfluoroalkyl substances (PFASs)? *Environmental Science & Technology*, 51(5), 2508-2518. https://doi.org/10.1021/acs.est.6b04806