Distribution Agreement

Burdens of Per- And Polyfluoroalkyl Substances (PFAS) In U.S. Public Water Systems

By

Madison Gabriella Lee
Master of Public Health

Environmental Health

Matthew O. Gribble, PhD DABT
Committee Chair

Burdens of Per- And Polyfluoroalkyl Substances (PFAS) In U.S. Public Water Systems

By

Madison Gabriella Lee
B.S.
North Carolina State University
2019

Thesis Committee Chair: Matthew O. Gribble, PhD DABT

An abstract of
a thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in Environmental Health
2021

Abstract

Burdens of Per- And Polyfluoroalkyl Substances (PFAS) In U.S. Public Water Systems

By Madison Gabriella Lee

Objective: Identify relationships between the detection of per- and polyfluoroalkyl substances (PFAS) in public water systems and county-level sociodemographic characteristics across the United States.

Methods: Drinking water and concentrations of six per- and polyfluoroalkyl substances were obtained from the United States Environmental Protection Agency's (U.S. EPA) Third Unregulated Contaminant Monitoring Rule (UCMR3), and county-level characteristics were obtained from the 2010 U.S. Census. The data from these two sources were used to construct classification trees to identify predictors of PFAS detection.

Results: The detection frequency of the six PFAS ranged between 0.05% and 1.00%. With these low detection frequencies, only PFHxS, PFOA, and PFOS data produced classification trees. The main predictors for the PFHxS, PFOA, and PFOS included different measures of household income, facility water type, population size, and residential mobility. Surface water as the facility water type was a common split among all three of the contaminants.

Conclusions: The classification trees were a novel approach to identifying disparities in the detection of PFAS in drinking water; however, the low detection frequencies from the 2012 – 2015 data limited potential subgroups of importance. The same approach should be used in future PFAS drinking water data to better predict which demographic characteristics predict PFAS burden.

Keywords: per- and polyfluoroalkyl substances; PFAS; drinking water; environmental health disparities; classification trees

Burdens of Per- And Polyfluoroalkyl Substances (PFAS) In U.S. Public Water Systems

By

Madison Gabriella Lee

B.S.

North Carolina State University

2019

Thesis Committee Chair: Matthew O. Gribble, PhD DABT

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Environmental Health

2021

Acknowledgements

Throughout the laborious process of cleaning and analyzing data, as well as writing, I have received an abundance of encouragement, guidance, and support from my advisor, colleagues, family, and friends. Such academic and emotional support have made this thesis possible.

I would like to thank my advisor, Dr. Matthew O. Gribble, for his guidance and support throughout this process. The statistical methods evolved several times prior to landing on the methods described in this thesis; Dr. Gribble has taught me a range of statistical methods through many Zoom calls, which have made my time in this Master of Public Health program all the more rewarding.

I would also like to thank PhD candidate Yachen Zhu and Dr. Scott Bartell from the University of California, Irvine, and Daniell Toth from the Bureau of Labor Statistics for their guidance in debugging and writing the code.

In addition, I would like to thank my family and friends for their emotional support throughout this process. The encouragement to work through this thesis was largely supplied by their coffee and phone calls.

All of these individuals have helped me immensely in finishing this thesis and Master of Public Health program.

TABLE OF CONTENTS

1. INTRODUCTION

Poly- and perfluoroalkyl substances (PFAS) are a group of chemicals from commercial and industrial processes found to be ubiquitous in the environment.[1,2] Prior studies on PFAS contamination have largely focused on distribution in specific regions or states based on proximity to point sources (i.e. commercial and industrial sites) rather than nationwide distribution. For instance, a study focused on New Hampshire community drinking water near a former U.S. Air Force base concluded that PFHxS, PFOA, and PFOS were significantly higher in the water samples and serum samples of the community compared to NHANES 2013 – 2014 data.[3] Hu et al. found that to be indicative not just in New Hampshire but nationally; the number of industrial sites, military fire training areas, and wastewater treatment plants were significant predictors of PFAS detection in community drinking water.[1]

Human exposure to PFAS has also been established in the general population.[4] The major exposure pathways to PFAS include ingestion and inhalation; ingestion of contaminated drinking water and food, and inhalation of contaminated air.[2] Serum samples from participants of NHANES revealed serum concentrations of PFHxS, PFNA, PFOA, and PFOS in over 98% of samples.[4] Human exposure was then found to be associated with several human health effects, such as increased risk of thyroid disease, increased risk of decreased fertility, and increased risk of kidney and testicular cancer.[5]

Even though research has established PFAS exposure in the general population, there are limited studies investigating potential disparities in its presence in drinking water. The Study of Women's Health Across the Nations (SWAN) concluded that site and race/ethnicity were significant predictors of PFAS; however, their study was limited to women aged 45-56 years in seven cities.[6] Additional research is needed with a wider scope of demographics to fully elucidate the distribution of PFAS in community drinking water on a national scale and potentially identify vulnerable populations. This study used publicly available drinking water data and demographic variables on the county-level to address any burdens of six perfluorinated compounds, including perfluorobutanesulfonic acid (PFBS), perfluoroheptanoic acid (PFHpA), perfluorohexanesulphonic acid (PFHxS), perfluorononanoic acid (PFNA), perfluorooctanoic acid (PFOA), and perfluorooctanesulfonic acid (PFOS).

2. MATERIALS AND METHODS

2.1 Data Sources

Public water system (PWS) data was obtained from the Third Unregulated Contaminant Monitoring Rule (UCMR3) in the U.S. EPA's National Occurrence Database (NCOD). The UCMR3 data provided information about public water systems and monitoring of 30 contaminants between January 2013 and December 2015. All large water systems, defined as serving more than 10,000 customers, were required to monitor the contaminants, in addition to a representative sample of small water systems, defined as serving 10,000 or fewer customers.

Only public water systems within the 50 states. Other relevant characteristics from the UCMR3 data included the PWS ID, facility size, facility water type, state, and zip code. The PWS ID, state, and zip code were used to retrieve the county served by the water system. A total of 1,878 counties were represented.

County demographics were obtained from the 2010 data by the U.S. Census Bureau. The demographics were assumed to be constant between census data collection and UCMR3 sampling. Demographics included variables on age and sex; education; businesses; computer and internet use; economy; families and living arrangements; health; housing; income and poverty; population size; race and Hispanic origin; and transportation.

2.2 Perfluorinated Contaminants

This analysis was restricted to the six perfluorinated contaminants included in UCMR3. Three of the perfluorinated contaminants were perfluoroalkyl carboxylic acids (PFCAs): PFHpA, PFNA, and PFOA. The other three were perfluoroalkane sulfonates (PFSAs): PFBS, PFHxS, and PFOS. Each of the contaminants were analyzed individually. Only 0.52% of the samples had concentrations of perfluorinated contaminants above the minimum reporting limit (MRL) using EPA Method 537 (0.05% for PFBS; 0.56% for PFHpA; 0.56% for PFHxS; 0.05% for PFNA; 1.00% for PFOA and 0.79% for PFOS). Therefore, a binary variable was created to assess whether the public water system had ever detected the individual contaminant between 2013 and 2015.

2.3 Statistical Analyses

Classification trees were used to assess which county characteristics were predictive of detected contaminants. Statistical analyses were conducted in R version 4.0.2. The goal was to predict which public water systems would have detectable levels of each contaminant based on the county demographic variables. The *rpart* program uses recursive partitioning for classification, regression, and survival trees. The "method" was specified as "class" to building classification models for each of the contaminants. Additional parameters required that the minimum number of observations in a node prior to splitting be 30 (minsplit = 30) and the complexity factor by 0.001 (cp = 0.001). With the complexity factor set at 0.001, a split will not be made if it does not increase the fit by a factor of 0.001.

The trees were built through splitting the data into sub-groups. Each sub-group was further split until no improvement could be made to the sub-group and/or the sub-group reached the minimum sample size.

Each of the county demographic variables were continuous; the program assessed the optimal split point for each of these variables.

3. RESULTS

Each of the perfluorinated contaminants were detected with a frequency ranging between 0.05% and 1.00%. PFBS and PFNA had the lowest detection frequencies at 0.05%, while PFOA had the highest detection frequency at 1.00%. There were 379 total samples with detected PFOA representing 117 PWSs. The summary of each perfluorinated contaminant are presented in **Table 1**.

The classification trees were created to predict the relationships between detectable compounds and county-level sociodemographic characteristics. However, no subgroups were created for PFBS, PFHpA, and PFNA due to the low detection frequencies and variation in sociodemographic characteristics. Therefore, only the classification trees for PFHxS, PFOA, and PFOS are presented. The classification trees for PFHxS is displayed in **Figure 1**.

Living in the same house as one year ago was the root node (Partition I) for PFHxS detection. The public water systems with the highest predicted probability of detection of PFHxS (0.83) were those with less than 78% of the population living in the same house as one year ago (Partition I), population greater than or equal to 483,700 (Partition II), and not solely using surface water as the source for their facility (Partition III). There were 0.7% observations in that end node, or 54 out of 7,760 public water systems. The classification tree for PFOA is displayed in **Figure 2**.

Monthly owner costs without mortgage was the root node (Partition I) for PFOA. The public water systems with the highest predicted probability of detection of PFOA (0.76) were those with monthly owner costs without mortgage less than $1,002 (Partition I), median value of owner-occupied housing units between $254,100 and $254,800 (Partition II and III), and not solely using surface water as the source for their facility (Partition IV). There were 0.6% of observations in the end node, or 48 out of 8,044 public water systems. The classification tree for PFOS is displayed in **Figure 3**.

The number of veterans was the root node (Partition I) for PFOS. The public water systems with the highest predicted probability of detection of PFOS (0.79) were those with greater than or equal to 83,930 veterans (Partition I), greater than or equal to 67% white (Partition II), and not solely using surface water as the source for their facility (Partition III). There were 0.48% of observations in the end node, or 37 out of 7,867 public water systems.

4. DISCUSSION

Given the detection frequencies for PFBS and PFNA at 0.05%, classification trees could not be produced. As presented in **Table 1**, PFBS was only detected in 8 PWSs and PFNA was only detected in 14 PWSs. However, the geometric mean concentration of PFNA in those 14 PWSs was 0.034 ppb, which exceeds the Minimum Risk Levels (MRLs) for children (0.021 ppb) set by the Agency for Toxic Substances and Disease Registry (ATSDR).[7] Additionally, the geometric mean concentrations of detected PFOA and PFOS exceeded MRLs. The geometric mean concentration for PFOA was 0.031 ppb, exceeding the MRL for children (0.021 ppb); the geometric mean concentration for PFOS was 0.081 ppb, exceeding the MRL for adults (0.052 ppb) and children (0.014 ppb). Therefore, while the detectable concentrations represent a small sample of public water systems, these exceedances require further evaluation and public health notice.

There was also no classification tree produced for PFHpA; despite the detection frequency (0.56%), there were no optimal subgroups found to be predictive of detection or non-detection. However, the data for PFHxS, PFOA, and PFOS were all successfully ran with the *rpart* program. Living in the same house as one year ago was an important overall predictor for the detection of PFHxS. It should be noted that the characteristic of living in the same house as one year ago is measured by the U.S. Census Bureau to assess residential stability and to understand the extent of residential migration and mobility. Higher residential mobility rates are higher among low-income households.[8] Additional factors contributing to higher residential mobility are household characteristics, housing unit conditions, metropolitan area and housing market dynamics, neighborhood dynamics, etc.[8] The optimal split for living in the same house as one year ago was 78%; the end node with the highest predicted probability for detection of PFHxS were public water systems with greater than 78% living in the same house as one year ago. In addition to residence on year ago, facility water type and population size were also important predictors for detection of PFHxS.

For PFOA, median selected monthly owner costs for housing units without a mortgage was an important overall predictor. The other important predictors were the median value of owner-occupied housing units and facility water type. Similarly, the important predictors for PFOS were veteran status, median value of owner-occupied housing units, population size, and facility water type. Given the option for veterans to reside on military bases and/or reside in the neighboring area of military bases, veteran status could be indicative of military sites which Hu et al. concluded to be predictive of PFAS detection in community drinking water.[1] However, additional data is needed to fully elucidate the link between the two.

The classification trees for PFHxS, PFOA, and PFOS all included facility water type. The UCMR3 data provided four different designations for water type: SW (surface water), GW (ground water), GU (ground water under the direct influence of surface water), and MX (any combination of SW, GW, and GU).

PFHxS and PFOS both split based on surface water, while PFOA had two splits: one based on surface water and one based on ground water. Much of the literature surrounding point sources of PFAS and drinking water support the finding of surface water being a predictor of PFHxS, PFOA, and PFOS detection.[9,10] Additionally, ground water has been subject to accumulation of PFAS containing aqueous film-forming foams (AFFF) used at firefighting training sites and military bases.[2] AFFF contamination in drinking water was identified as a nationally significant challenge in the United States.[2]

In contrast to Park et al., this study did not identify ethnicity or race to be major predictors of all PFAS chemicals analyzed.[6] Only the classification tree for PFOS identified that greater than or equal to 67.15% white (not Hispanic) was predictive of detection. However, the classification tree for PFOA identified monthly owner costs without mortgage and median value of owner-occupied housing units as predictive of detection. Those two predictors are proxies to household income and overall socio-economic status (SES). While SES has several contributing economic and educational patterns, it is also related to race.[11] However, it should be noted that SES and race are not interchangeable; Williams et al. notes that race is significant in addressing health disparities because SES indicators are not equivalent across racial groups.[11]

4.1 Limitations

There were several limitations to acknowledge in this study. There were low detection frequencies of all perfluorinated contaminants in this analysis, which could be attributed to the detection limit of the analytical techniques throughout the analysis period (2013 – 2015) and/or to low concentrations of the perfluorinated contaminants in the public drinking water systems. The lowest minimum reporting limit during the collection of UCMR3 data was 0.010 ppb and has seen been improved to as low as 0.001 ppb. This study also only analyzes the data provided from the public water systems; the majority of exposure of PFAS has been linked to oral ingestion, including fish and shellfish, however, only drinking water is considered here.[12,13] Sunderland et al. noted that exposure to PFAS from drinking water and shellfish are increasing or stable in several regions.[2]

Additionally, the analyses assume that the sociodemographic characteristics provided by the 2010 Census remained constant throughout the sampling period. Even though a single county may be served by more than one public water system, the analyses assume that the county is homogenous, and its demographics are representative of each of the individuals served by public water systems.

The UCMR3 data also excludes private drinking water systems or wells that are also at risk for PFAS contamination. The UCMR3 data also only addressed six perfluorinated compounds; however, there are several more emerging PFAS threatening safe drinking water.

5. PUBLIC HEALTH IMPLICATIONS

The detection of PFAS in public drinking water systems given the analytical methods between 2013 and 2015 is significant for public health; of the public water systems with detectable concentrations, the geometric means for PFNA, PFOA, and PFOS exceeded Minimum Risk Levels sets by the Agency for Toxic Substances and Disease Registry. Additional attention is needed for those public water systems.

The classification tree analysis was a novel approach for disparities research than can be used with data using analytical methods with lower minimum reporting limits for PFAS. With potentially higher detection frequencies among the perfluorinated compounds more subgroups may be identified to better understand the disparities in drinking water.

REFERENCES

(1) Hu, X. C.; Andrews, D. Q.; Lindstrom, A. B.; Bruton, T. A.; Schaider, L. A.; Grandjean, P.; Lohmann, R.; Carignan, C. C.; Blum, A.; Balan, S. A.; Higgins, C. P.; Sunderland, E. M. Detection of Poly- and Perfluoroalkyl Substances (PFASs) in U.S. Drinking Water Linked to Industrial Sites, Military Fire Training Areas, and Wastewater Treatment Plants. *Environ. Sci. Technol. Lett.* **2016**, *3* (10), 344–350. https://doi.org/10.1021/acs.estlett.6b00260.

(2) Sunderland, E. M.; Hu, X. C.; Dassuncao, C.; Tokranov, A. K.; Wagner, C. C.; Allen, J. G. A Review of the Pathways of Human Exposure to Poly- and Perfluoroalkyl Substances (PFASs) and Present Understanding of Health Effects. *J Expo Sci Environ Epidemiol* **2019**, *29* (2), 131–147. https://doi.org/10.1038/s41370-018-0094-1.

(3) Daly, E. R.; Chan, B. P.; Talbot, E. A.; Nassif, J.; Bean, C.; Cavallo, S. J.; Metcalf, E.; Simone, K.; Woolf, A. D. Per- and Polyfluoroalkyl Substance (PFAS) Exposure Assessment in a Community Exposed to Contaminated Drinking Water, New Hampshire, 2015. *International Journal of Hygiene and Environmental Health* **2018**, *221* (3), 569–577. https://doi.org/10.1016/j.ijheh.2018.02.007.

(4) Calafat, A. M.; Wong, L.-Y.; Kuklenyik, Z.; Reidy, J. A.; Needham, L. L. Polyfluoroalkyl Chemicals in the U.S. Population: Data from the National Health and Nutrition Examination Survey (NHANES) 2003–2004 and Comparisons with NHANES 1999–2000. *Environ Health Perspect* **2007**, *115* (11), 1596–1602. https://doi.org/10.1289/ehp.10598.

(5) Calafat, A. M.; Kato, K.; Hubbard, K.; Jia, T.; Botelho, J. C.; Wong, L.-Y. Legacy and Alternative Per- and Polyfluoroalkyl Substances in the U.S. General Population: Paired Serum-Urine Data from the 2013–2014 National Health and Nutrition Examination Survey. *Environment International* **2019**, *131*, 105048. https://doi.org/10.1016/j.envint.2019.105048.

(6) Park, S. K.; Peng, Q.; Ding, N.; Mukherjee, B.; Harlow, S. D. Determinants of Per- and Polyfluoroalkyl Substances (PFAS) in Midlife Women: Evidence of Racial/Ethnic and Geographic Differences in PFAS Exposure. *Environmental Research* **2019**, *175*, 186–199. https://doi.org/10.1016/j.envres.2019.05.028.

(7) PFAS Minimal Risk Levels and Environmental Media Evaluation Guides https://www.atsdr.cdc.gov/pfas/resources/mrl-pfas.html (accessed Mar 29, 2021).

(8) Family Residential Instability: What Can States and Localities Do? 24.

(9) Banzhaf, S.; Filipovic, M.; Lewis, J.; Sparrenbom, C. J.; Barthel, R. A Review of Contamination of Surface-, Ground-, and Drinking Water in Sweden by Perfluoroalkyl and Polyfluoroalkyl Substances (PFASs). *Ambio* **2017**, *46* (3), 335–346. https://doi.org/10.1007/s13280-016-0848-8.

(10) Nguyen, T.; Reinhard, M.; Chen, H.; Gin, K. Fate and Transport of Perfluoro- and Polyfluoroalkyl Substances Including Perfluorooctane Sulfonamides in a Managed Urban Water Body. *Environmental Science & Pollution Research* **2016**, *23* (11), 10382–10392. https://doi.org/10.1007/s11356-016-6788-9.

(11) Williams, D. R.; Priest, N.; Anderson, N. Understanding Associations between Race, Socioeconomic Status and Health: Patterns and Prospects. *Health Psychol* **2016**, *35* (4), 407–411. https://doi.org/10.1037/hea0000242.

(12) Christensen, K. Y.; Raymond, M.; Blackowicz, M.; Liu, Y.; Thompson, B. A.; Anderson, H. A.; Turyk, M. Perfluoroalkyl Substances and Fish Consumption. *Environmental Research* **2017**, *154*, 145–151. https://doi.org/10.1016/j.envres.2016.12.032.

(13) Fromme, H.; Tittlemier, S. A.; Völkel, W.; Wilhelm, M.; Twardella, D. Perfluorinated Compounds – Exposure Assessment for the General Population in Western Countries. *International Journal of Hygiene and Environmental Health* **2009**, *212* (3), 239–270. https://doi.org/10.1016/j.ijheh.2008.04.007.

TABLES AND FIGURES

Table 1.

Summary Statistics of Perfluorinated Contaminants in Public Water Systems (PWSs)

Contaminant	Detection Frequency, n (%)		PWSs with Detect	parts per billion (ppb) Minimum Reporting Limit (MRL)	Geometric Mean (Standard Deviation)	
PFBS	19	(0.05)	8	0.09	0.154	(1.453)
PFHpA	236	(0.56)	86	0.01	0.012	(1.728)
PFHxS	207	(0.56)	55	0.03	0.092	(2.209)
PFNA	19	(0.05)	14	0.02	0.034	(1.337)
PFOA	379	(1.00)	117	0.02	0.031	(1.600)
PFOS	292	(0.79)	95	0.02	0.081	(2.164)

Fig 1

Classification Tree for Detection of PFHxS. Classification tree for detection of PFHxS with each end node specifying non-detection (0) or detection (1), predicted probability, and percent of observations in the end node. Each binary split labeled by Roman numerals are detailed in **Table 2**.

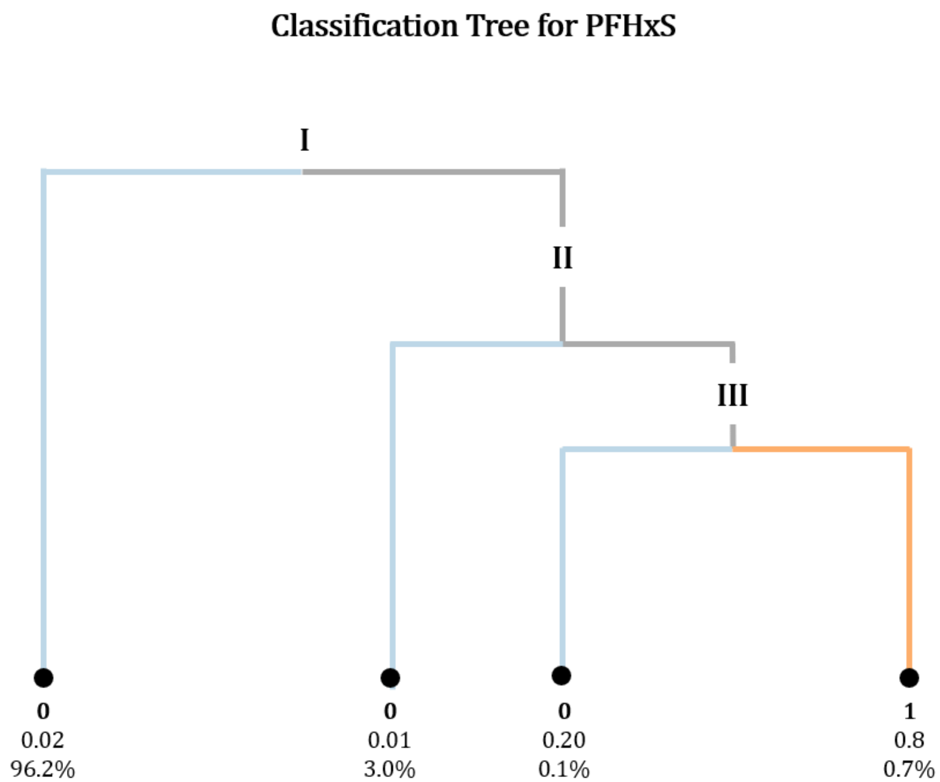

Table 2.

Detection of PFHxS for Each Node in the Classification Tree

Partition	Node	End Node	Split	Predicted Class	Predicted Probability	Observations in Node (%)
I	Root	No	Living in Same House 1 Year Ago < 77.55%	0	0.03	100.0
I	L	Yes	Living in Same House 1 Year Ago ≥ 77.55%	0	0.02	96.2
II	L	Yes	Population < 483,700	0	0.01	3.0
II	R	No	Population ≥ 483,700	1	0.73	0.8
III	L	Yes	Facility Water Type = Surface Water	0	0.20	0.1
III	R	Yes	Facility Water Type ≠ Surface Water*	1	0.83	0.7

Node – L: Left Node, R: Right Node

Predicted Class – 0: No detection of PFHxS, 1: Detection of PFHxS

* Other facility water types include ground water, ground water under the direct influence of surface water, and a combination of surface water and ground water

Fig 2.

Classification Tree for Detection of PFOA. Classification tree for detection of PFOA with each end node specifying non-detection (0) or detection (1), predicted probability, and percent of observations in the end node. Each binary split labeled by Roman numerals are detailed in **Table 3**.

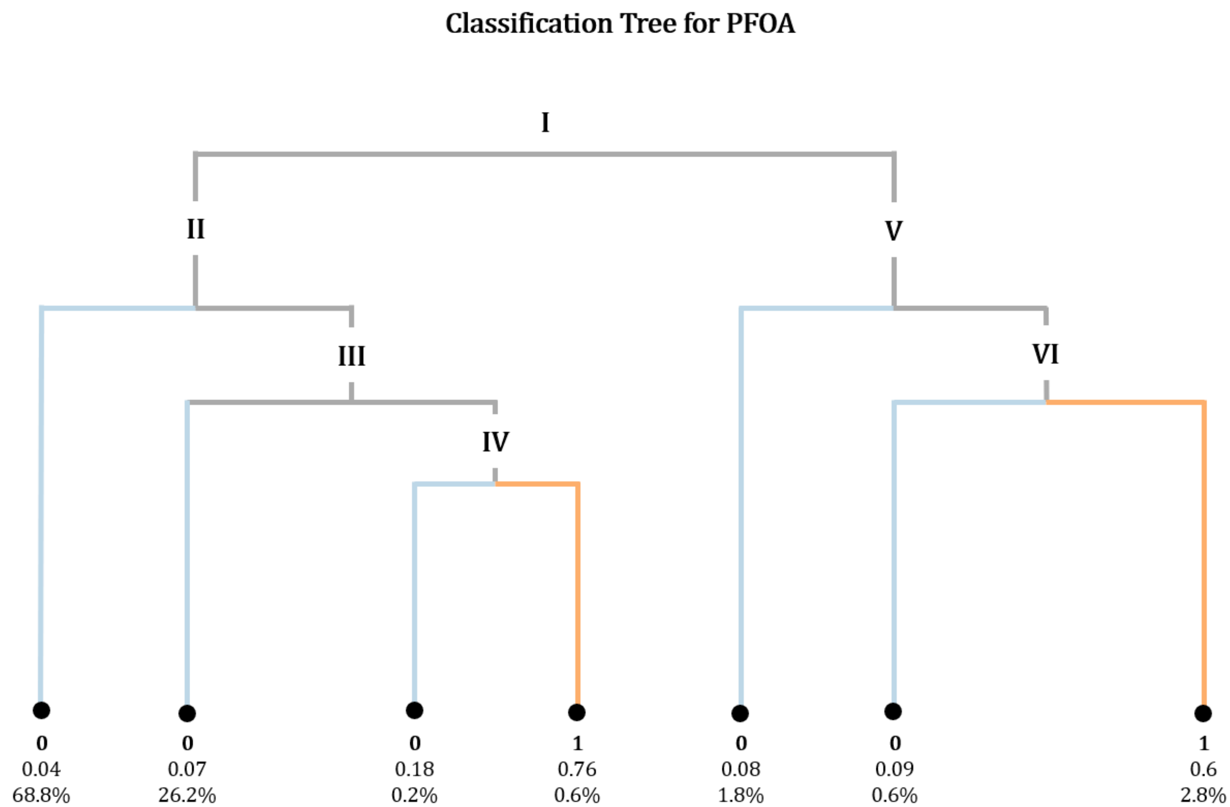

Table 3.

Detection of PFOA for Each Node in the Classification Tree

Partition	Node	End Node	Split	Predicted Class	Predicted Probability	Observations in Node (%)
I	Root	No	Monthly Owner Costs without Mortgage < $1,002	0	0.07	100.0
II	L	Yes	Median Value of Owner-Occupied Housing Units < $ 254,100	0	0.04	67.8
II	R	No	Median Value of Owner-Occupied Housing Units ≥ $ 254,100	0	0.08	27.0
III	L	Yes	Median Value of Owner-Occupied Housing Units ≥ $ 254,800	0	0.07	26.2
III	R	No	Median Value of Owner-Occupied Housing Units < $ 254,800	1	0.61	0.8
IV	L	Yes	Facility Water Type = Surface Water	0	0.18	0.2
IV	R	Yes	Facility Water Type ≠ Surface Water*	1	0.76	0.6
V	L	Yes	Facility Water Type = Ground Water	0	0.08	1.8
V	R	No	Facility Water Type ≠ Ground Water	1	0.54	3.4
VI	L	Yes	Population ≥ 911,000	0	0.09	0.6
VI	R	Yes	Population < 911,000	1	0.64	2.8

Node – L: Left Node, R: Right Node

Predicted Class – 0: No detection of PFOA, 1: Detection of PFOA

* Other facility water types include ground water, ground water under the direct influence of surface water, and a combination of surface water and ground water

Fig 3.

Classification Tree for Detection of PFOS. Classification tree for detection of PFOS with each end node specifying non-detection (0) or detection (1), predicted probability, and percent of observations in the end node. Each binary split labeled by Roman numerals are detailed in **Table 4**.

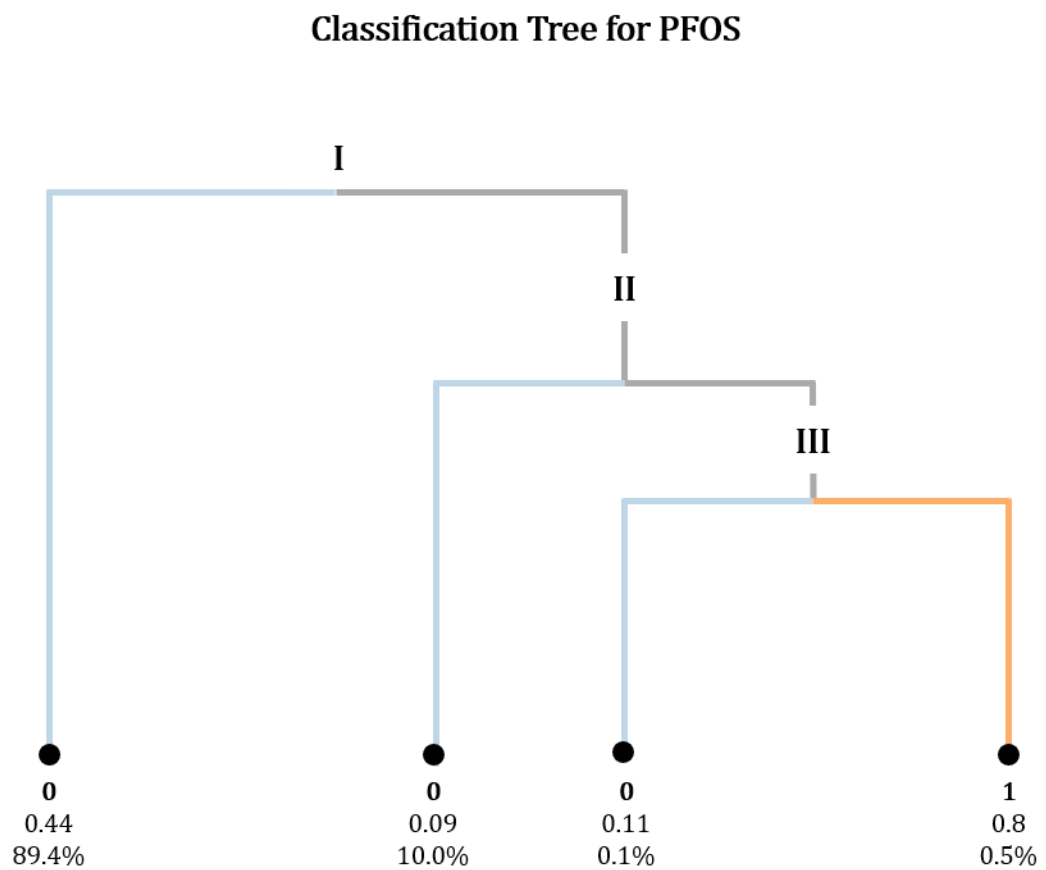

Table 4.

Detection of PFOS for Each Node in the Classification Tree

Partition	Node	End Node	Split	Predicted Class	Predicted Probability	Observations in Node (%)
I	Root	No	Veteran Count < 83,930	0	0.04	100.0
I	L	Yes	Veteran Count < 83,930	0	0.12	10.6
II	L	Yes	White (Not Hispanic) < 67.15%	0	0.09	10.0
II	R	No	White (Not Hispanic) ≥ 67.15%	1	0.66	0.6
III	L	Yes	Facility Water Type = Surface Water	0	0.11	0.1
III	R	Yes	Facility Water Type ≠ Surface Water*	1	0.79	0.48

Node – L: Left Node, R: Right Node
Predicted Class – 0: No detection of PFOS, 1: Detection of PFOS
* Other facility water types include ground water, ground water under the direct influence of surface water, and a combination of surface water and ground water