McAdam *et al. Water Emerg Contam Nanoplastics* 2024;3:15
DOI: 10.20517/wecn.2024.05

Water Emerging Contaminants & Nanoplastics

Research Article

Open Access

Check for updates

Exogenous factors associated with legacy PFAS concentrations in the general U.S. population: NHANES 1999-2018

Jordan McAdam[1], Laura E. Jones[2,3], Xiaobo Xue Romeiko[1], Erin M. Bell[1,3]

[1]Department of Environmental Health Sciences, University at Albany, State University of New York, Rensselaer, NY12144, USA.
[2]Center for Biostatistics, Bassett Research Institute, Cooperstown, NY 13326, USA.
[3]Department of Epidemiology and Biostatistics, School of Public Health, State University of New York, Rensselaer, NY 12144, USA.

Correspondence to: Dr. Erin M. Bell, Departments of Epidemiology and Biostatistics, School of Public Health, State University of New York, 1 University Place, Rensselaer, NY 12144, USA. E-mail:ebell@albany.edu

How to cite this article: McAdam J, Jones LE, Romeiko XX, Bell EM. Exogenous factors associated with legacy PFAS concentrations in the general U.S. population: NHANES 1999-2018. *Water Emerg Contam Nanoplastics* 2024;3:15. https://dx.doi.org/10.20517/wecn.2024.05

Received: 7 Mar 2024 **First Decision:** 26 Apr 2024 **Revised:** 21 May 2024 **Accepted:** 23 May 2024 **Published:** 31 May 2024

Academic Editor: Joana C Prata **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

To reduce per- and polyfluoroalkyl substances (PFAS) exposure through legislation and other interventions, we must understand factors contributing to individual body burden. Identifying factors associated with serum perfluorinated compounds (PFAS) concentrations also aids in identifying groups at higher risk of adverse health outcomes due to elevated exposure. This exploratory analysis provides initial findings on exposure-related factors associated with legacy PFAS concentrations in the general United States (U.S.) population. We obtained National Health and Nutrition Examination Survey (NHANES) datasets with individual serum PFAS measurements from cycle years 1999-2018 (N = 14,961), excluding the 2001-2002 cycle due to pooled sampling. Over 100 features were evaluated for associations with PFAS concentrations. Data were singly imputed using hot deck and predictive mean matching (PMM) methods, and model performance by imputation method was compared using elastic net regression (ENR) models. Hot-deck imputation explained the most variance in predictive models; thus, a dataset imputed via hot-deck was used for feature selection. ENR models were employed to identify the top variables associated with legacy PFAS concentrations, and selected features were put into linear mixed models to obtain beta estimates and standard errors. Survey year, demographic, income, place of birth and citizenship status,

household size, dietary, health, food insecurity, general health and healthcare, housing, social and behavioral, and other characteristics were important factors associated with legacy PFAS concentrations in this nationally representative study population. A better understanding of exogenous factors associated with PFAS concentrations can influence future epidemiological studies, guiding decisions on adjustment for confounding, and advancing our understanding of factors that affect chemical half-lives and toxicokinetics.

Keywords: PFAS, predictors, determinants, demographics, consumer behavior, healthcare, socioeconomics

INTRODUCTION

Per- and polyfluoroalkyl substances (PFAS) are persistent organic pollutants (POPs) and up to 99% of the United States (U.S.) human population has detectable concentrations of PFAS in their blood[1,2]. Though two of the most studied PFAS [i.e., perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA)] were phased out of production in many developed countries, concern over health effects associated with exposure to PFOA, PFOS, and other PFAS continues due to their environmental persistence. Many adverse health effects have been associated with PFAS exposure in adults, including dyslipidemia[3-5], changes in liver enzymes[6,7], elevated blood pressure and hypertension[8,9], and kidney and testicular cancer[10]. Because PFAS bioaccumulate in the bodies of adults, transplacental transfer, *in utero*, and breastfeeding, postnatally, can be sources of exposure in offspring[11].

Human exposure to PFAS is complex and varies within and between populations; however, ingestion has been suggested as the main route of exposure to PFAS among the general population. Foods such as fish, meat, offal, eggs, and fruit can be environmentally contaminated[12], but PFAS can also migrate into food from packaging[13-16]. Coated cookware can also leach PFAS into food[17]. Some communities in the U.S. have been highly exposed to PFAS through ingestion of contaminated drinking water[17], though this is not considered a major route of exposure for the general U.S. population. Other routes of PFAS exposure include ingestion of dust, transfer from treated carpets, clothes, and upholstery, inhalation of air, and inhalation of impregnation spray aerosols[18,19]. Though not thought to be a primary route of exposure, there may be dermal exposure from wearing clothing, such as waterproof layers, and from personal care products[18,19]. Serum or plasma concentrations of PFAS are generally used as markers of PFAS body burdens in humans, as they reflect long-term cumulative exposure to the chemicals from all sources of exposure.

To reduce PFAS exposure through legislation and other public health interventions, it is critical to understand factors that may contribute to individual body burden. Identifying factors associated with serum PFAS concentrations can also aid in identifying groups at higher risk of adverse health outcomes due to elevated exposure. Some factors associated with PFAS concentrations, such as parity and breastfeeding history, have been better studied[20-22], but less is known about other factors that may contribute to PFAS body burden in the general adult population. Previous studies evaluating factors associated with PFAS concentrations analyzed a limited number of demographic or dietary factors[23-26]. In addition, many have been conducted in European or Asian populations[24,25,27,28] and are not generalizable to the U.S. population. Given demographic, social and behavioral habits, and regulatory differences, it is important to consider the general U.S. population separately. There has also been a focus on populations living in PFAS-contaminated communities[26,28-30], sensitive populations (i.e., pregnant mothers and neonates)[31-37], and specific demographic or occupational groups[38-40], which are not generalizable to populations with lower, background levels of exposure and non-pregnant adult populations. Due to the paucity of research evaluating a large number of potential factors associated with serum PFAS concentrations in the general U.S. adult population, we aimed to investigate these factors using data from nine cycles of U.S. national health surveys [National Health and Nutrition Examination Survey (NHANES)] as an exploratory analysis.

Unlike previous analyses, we evaluate many potential factors associated with PFAS concentrations using the rich questionnaire data available from NHANES. In this analysis, we focus on exogenous factors (i.e., factors that may be assessed through a questionnaire such as social and behavioral habits, demographic, and household characteristics) that may be associated with PFAS exposure and serum concentrations, and in a separate analysis, we will explore endogenous factors (i.e., factors that may be measured through bloodwork or body measures such as complete blood count panels, environmental chemical concentrations, or body mass index) associated with PFAS concentrations. We evaluate these groups of factors separately due to their potential implications. A better understanding of the exogenous factors associated with PFAS concentrations may influence future epidemiological studies by guiding decisions on adjustment for confounding and furthering toxicological studies by advancing our understanding of factors that affect chemical half-lives, toxicokinetics, and interindividual variation[41,42]. This study is warranted, given that PFAS exposure is multifactorial and complex.

EXPERIMENTAL

Study population

The NHANES is a nationally representative survey conducted by the U.S. National Center for Health Statistics that is designed to assess the health and nutrition status of adults and children in the U.S. through in-home interviews and medical examinations[43]. In this study, we analyzed publicly available data from nine cycles of the NHANES (1999-2000, 2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, 2017-2018). We did not include data collected in the 2001-2002 cycle because PFAS were measured in pooled, rather than individual, serum samples during those years. All study protocols were approved by the National Center for Health Statistics institutional review board and all participants gave written informed consent.

We restricted analyses to participants aged 18 years and older with complete serum PFAS measurements, as there may be differences in social and behavioral factors contributing to PFAS exposure and body burden between adults and children and differences in health outcome etiology for children following PFAS exposure. The final analytic sample was comprised of 14,961 adults. A flow chart showing sample selection and the workflow pipeline is shown in Supplementary Figure 1.

Exposure assessment

Solid phase extraction-high performance liquid chromatography-turbo ion spray ionization-tandem mass spectrometry was used to measure PFAS concentrations in serum samples. Detailed descriptions of the analytic methods have been described previously[44,45]. Our analysis focused on four highly detected PFAS: PFOS, PFOA, perfluorohexane sulfonate (PFHxS), and perfluorononanoic acid (PFNA). Non-detectable concentrations of PFAS were substituted with the respective limit of detection divided by the square root of two[46]. In most cycles, concentrations of PFOS and PFOA were measured in total; however, the 2013-2014, 2015-2016, and 2017-2018 cycles measured linear and branched isomers. To maximize our study population and because many of the regulations and much of the medical guidance surrounding PFAS exposure are based on the summed measures of these congeners[47], we summed these isomers to calculate total PFOS and PFOA concentrations. In the 1999-2000 cycle, PFAS were measured only among individuals with enough surplus sera (approximately 17.8% of participants). For subsequent cycles, PFAS were measured in approximately one-third of the sub-sample of the study cohort at random.

Evaluation of factors associated with serum PFAS concentrations

Each survey cycle contains five components: demographics, dietary, examination, laboratory, and questionnaire data. Each data component has many data files depending on the information collected in that round's survey and examination assessment. These files are in SAS format and available for download

at https://wwwn.cdc.gov/nchs/nhanes/Default.aspx. Datasets used in this analysis from NHANES were downloaded in October and November 2022. The five data components are described below:

- Demographic data: provide individual, family, and household-level information on language, pregnancy status, income, family size, household composition, demographic information, and other selected information such as military service status and country of birth.
- Dietary data: provide detailed dietary intake information from participants, including a 24-hour food and beverage recall, nutrient intake estimates, and supplement use.
- Laboratory and examination data: includes physical measures that can include blood pressure, body measures, oral health examinations, and physical activity monitoring.
- Questionnaire data: provide information on behavior, diet, nutrition, and other factors.

This analysis focused only on potential exogenous factors associated with PFAS concentrations, as described above; therefore, only questionnaire data were utilized in addition to the laboratory measurements of serum PFAS concentrations. We considered data over all cycles combined rather than single cycles at a time to increase the sample size[48].

Given differences across cycles, some variables had a high percent missingness in this larger, combined dataset with all 14,961 participants. Missingness ranged from 0% to 99.38% in the dataset, and 575 variables were identified across all cycles. As PFAS exposure has been associated with numerous health outcomes, variables related to diabetes, blood pressure, kidney conditions, asthma, arthritis, thyroid problems, liver conditions, and cancer, as well as medication use related to these conditions, were excluded from this analysis, given the evidence that PFAS exposure may contribute to these conditions[10,49-55]. Those variables with more than 50% missingness were excluded. This left 100 variables plus 14 demographic covariates for analysis [Figure 1]. The exogenous variables assessed as potential factors associated with serum PFAS concentrations are detailed in Supplementary Table 1. Certain variables from NHANES were re-coded for use in these analyses due to changes in question format across cycles, which are described in Supplementary Table 1. Unless otherwise noted, variables were not re-categorized from what is described in NHANES documentation (https://wwwn.cdc.gov/nchs/nhanes/default.aspx). Reference categories are described where applicable.

Statistical analysis

Data distributions and ranges were calculated and univariate analysis, including heatmaps and intersection analysis to explore missingness between pairs of features, was performed to assess the structure of missingness in the data. First, we describe imputation methods employed to address missingness. Next, we describe methods used to compare imputation methods and select an imputation method for model building. We then describe methods to identify variables best explaining variance in PFAS concentrations and analyses by congener.

Imputation

Because this was a predictive study employing large datasets and we were not performing inference, multiple imputation (MI) - which can be slow and requires pooling of statistical results for inference - was not necessary and was difficult to manage in a machine-learning context[56]. In addition, traditional MI methods have performed poorly on high-dimensional data[57]. We thus chose to explore two efficient donor-based single imputation methods: hot deck and predictive mean matching (PMM).

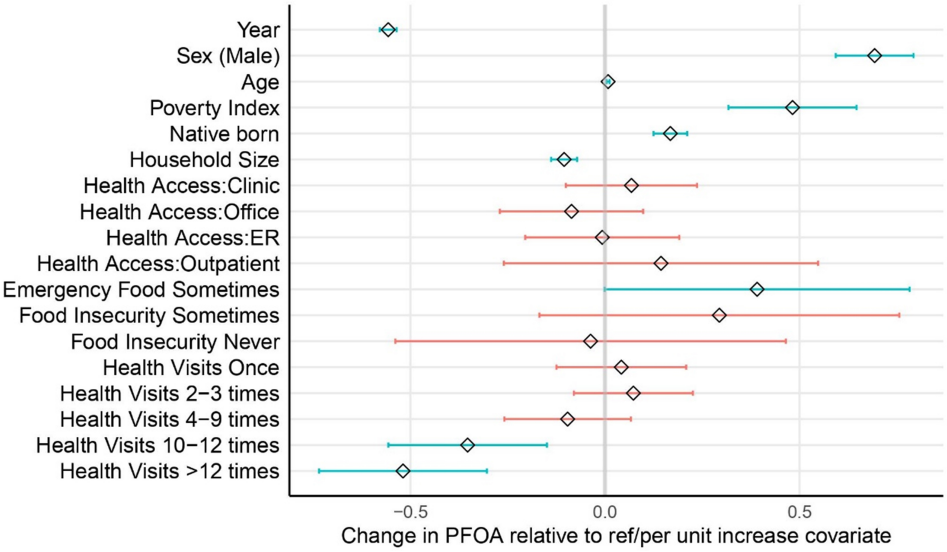

Figure 1. Change in serum PFOA relative to the reference level (for categorical variables) or per unit SD increase for continuous covariates from multivariable mixed-effects regression. See Supplementary Table 2 for details. Blue indicates statistically significant findings, while orange indicates findings that are not statistically significant. PFOA: Perfluorooctanoic acid.

Up to 50% of samples were monotonically missing (i.e., with monotone missingness patterns) for subsets of covariates. Under the assumption that our data were missing at random (MAR), we performed imputation by grouped and sorted simple hot deck and by PMM as described below. Both are donor-based methods that work well with mixed continuous and categorical data and draw replacements from existing observations. Imputations are thus realistic, with no imputed values outside the observed range. Donor-recipient methods such as hot deck are fast and appropriate for large datasets[58].

Hot deck imputation

Simple hot deck is a fast and efficient donor-based method where replacements for missing units (recipients) are randomly selected from complete units (donors) in the same column. Since it is donor-based, hot deck imputation avoids issues with model misspecification that might arise in regression-based methods, and handles large-dimensioned datasets efficiently. While hot deck methods are extensively used, the theory behind the method is less well-developed than other imputation methods[59]. In Hot Deck methods, donor units are randomly selected from a set of complete cases, known as the donor pool. There are multiple ways to specify a donor pool, otherwise known as an imputation class, for each missing unit. For example, columns to be imputed can be sorted and grouped based on other metrics, and sampling occurs from complete units within groupings. We grouped by participant sex, sorted by age, and found appropriate donor values for each missing recipient within the sorted domains by sex and age. Data were imputed using the package "VIM" in R[60].

PMM

PMM is a semi-parametric hot deck method that handles both continuous and categorical data and is robust against model misspecification. As implemented in the "mice" package in R[61], PMM uses an imputation model, for example, regression, to predict missing values, and then matches the predicted value of the missing unit (recipient) with a set of candidate donors from all complete cases that have predicted values closest to that of the missing unit. A donor is randomly chosen from the candidates, and the observed value of the donor is taken to replace the missing value. We assumed donor sets of size five,

imputed five iterations, and then combined the datasets into one complete dataset.

Comparison of imputation methods

Elastic net regression (ENR) is a regularized regression method that includes the properties of both ridge regression and least absolute shrinkage and selection operator (LASSO)[62]. By linearly combining ridge and LASSO penalty functions, ENR performs variable selection and continuous shrinkage, and selects groups of correlated variables as a whole.

We compared imputation methods by building ENR models to assess model performance. Data preprocessing involved centering and scaling for continuous variables. All potential factors described in Supplementary Table 1 were entered into the ENR models and model metrics were obtained and compared.

When comparing the models with imputed data from each method, we considered the percent variance explained and R^2 metrics from the ENR model. The goal was to identify the imputation method with the best fitting models that explained the most variance in PFAS concentrations. The dataset from the best fitting imputation method was then used for identification of factors associated with PFAS concentrations.

Imputed datasets from each of the methods described above were randomly split 70/30 for training sets (used to train the models) and testing sets (used to test the models). This method was selected as it is appropriate for datasets with both a large number of rows and columns and is widely used for feature selection.

Identification of factors associated with serum PFAS concentrations

ENR models were used to identify variables that best explain variance in PFAS concentrations. We used the trainControl function in the caret package to specify parameters for model training, where we specified repeated 5-fold cross-validation to identify the optimal model parameters based on the smallest RMSE values. We then trained the ENR model based on the trainControl parameters and made predictions on the test dataset. The varImp (i.e., variable importance) function in the caret library was used to identify the 10 most important variables explaining variance in the five PFAS exposure variables [PFOA; PFOS; PFNA; PFHxS; and the simple sum of PFOA, PFOS, PFNA, and PFHxS (ΣPFAS)]. We computed the simple additive sum, rather than other approaches, to understand the body burden of mixtures of PFAS chemicals, to align with the current health guidance for PFAS exposure and testing that is based on a simple additive sum[47]. The variable importance represents the absolute value of ranked standardized coefficients from the best-fit tuned models. We did not conduct stratified analyses due to the exploratory nature of this analysis and due to differences in variables across the NHANES cycles included in this analysis.

Multivariable modeling

Based on results from feature selection described above, we built multivariable regression models for each of the five PFAS exposure variables, fitted with the 10 most important variables identified. Models were tuned by comparing tested models through analysis of variance (ANOVA) and using the lmeControl function. We extracted beta estimates and standard errors for model interpretation.

All data preprocessing, analysis, and validation were performed using R (version 3.4).

RESULTS AND DISCUSSION

Results

We present the demographic characteristics of participants in Table 1 (missingness was imputed via hot-deck). The smallest share of participants was from the 1999-2000 NHANES cycle (7.43%). The mean (standard deviation) age of participants in this analysis was 47.84 (19.24) years. Most participants were non-Hispanic (NH) White (39.87%-45.85%) and nearly one-third of participants had two persons in their household (29.80%). The PFAS with the highest median serum concentration was PFOS at 10.20 mg/dL. The median serum concentration of ΣPFAS was 16.40 mg/dL [Table 2].

After building the ENR models for each PFAS variable, statistics from the models were compiled to compare the results using the imputation method. Models based on grouped and sorted hot deck imputed data explained the greatest percent variance in PFAS concentrations for each PFAS and had the highest R^2 values for the PFOA, PFOS, PFHxS, and ΣPFAS models (data not shown). Given the overall better model statistics from hot deck imputation, we used this dataset for identification of important variables and for building linear mixed models. Overall, PFOA, PFOS, and ΣPFAS models had the best model fit, with R^2 values ranging from 0.21 to 0.31 (data not shown).

The top 10 variables explaining variance in serum PFAS concentrations, along with beta estimates and standard errors, are presented as forest plots in Figures 1-5 (see Supplementary Table 2 for a description of covariate codings on the Figures) and summarized in Table 3.

Male (relative to female) sex had the largest significant positive effects relative to all serum PFAS [Figures 1-5]. Year (i.e., NHANES cycle) has significant negative associations and had the largest (negative) effect size for PFOA [Figure 1] and PFOS [Figure 2], ΣPFAS [Figure 5] concentrations, suggesting diminishing population exposure or bioaccumulation in the sample populations over time. Age was significantly positively associated with all serum PFAS outcomes except for PFHxS [Figure 3], but effect sizes were small.

Place of birth and/or citizenship were important factors associated with serum PFOA [Figure 1], PFOS [Figure 2], PFHxS [Figure 3], and ΣPFAS concentrations, with being U.S.-born (relative to foreign-born) and possessing U.S. citizenship (relative to not being a U.S. citizen) associated with higher serum concentrations of these PFAS. Household size was an important factor in PFOA [Figure 1], PFOS [Figure 2], and ΣPFAS [Figure 5] models and was negatively associated with these PFAS concentrations. Race/ethnicity were also important factors associated with PFOS [Figure 2], PFNA [Figure 4], and ΣPFAS [Figure 5] concentrations, with the highest concentrations of these PFAS associated with NH Asian ethnicity and the lowest with Mexican-American race/ethnicity (reference group NH White). Three measures of fish and shellfish consumption were identified as important factors associated with serum PFAS concentrations in some models, with fish consumption positively and significantly associated with PFAS concentrations relative to no fish consumption. "Ever" *vs.* "never" freshwater fish consumption (in the last 30 days) was significantly positively associated with PFOS [Figure 2], PFNA [Figure 4], and ΣPFAS concentrations [Figure 5], while "ever" *vs.* "never" shellfish consumption (last 30 days) was significantly associated with PFOS [Figure 2], PFNA [Figure 4], and ΣPFAS [Figure 5] concentrations. "Ever" *vs.* "never" consumption of any fish in the last 30 days was an important factor associated with PFNA concentrations.

Several variables related to healthcare, income, and financial or food security were identified as important factors associated with serum PFAS concentrations. For the PFOA model, where healthcare is accessed (i.e., the ER, hospital outpatient department, a medical clinic, or private medical practice, relative to the reference "another place"), household food insecurity measures, and the number of times a participant received

Table 1. Demographic characteristics of participants, NHANES 1999-2018 (*N* = 14,961)

Variable	Level	*N* (%)
NHANES cycle	1999-2000	1,111 (7.43)
	2003-2004	1,680 (11.23)
	2005-2006	1,657 (11.08)
	2007-2008	1,834 (12.26)
	2009-2010	1,951 (13.04)
	2011-2012	1,648 (11.02)
	2013-2014	1,669 (11.16)
	2015-2016	1,711 (11.44)
	2017-2018	1,700 (11.36)
Sex (mean, SD)	Male	7,247 (48.44)
	Female	7,714 (51.56)
Age (mean, SD)	Years	47.84 ± 19.24
Veteran/military status	Yes	1,707 (11.41)
	No	13,252 (88.58)
Citizenship status	Citizen by birth	12,714 (84.98)
	Not a citizen	2,212 (14.79)
	Refused or don't know	35 (0.23)
Education level	Less than high school	4,228 (28.26)
	High school	3,629 (24.26)
	Greater than high school	7,081 (47.33)
	Refused or don't know	23 (0.15)
Marital status	Married	7,513 (50.22)
	Widowed	1,214 (8.11)
	Divorced	1,465 (9.79)
	Separated	478 (3.19)
	Never married	2,969 (19.84)
	Living with partner	1,316 (8.80)
	Refused	6 (0.04)
Race/ethnicity	Mexican-American	2,944 (19.68)
	Other Hispanic	836 (5.59)
	NH White	6,859 (45.85)
	NH Black	2,991 (19.99)
	Other race	476 (3.18)
	NH Asian	855 (5.71)
Total number of people in family	1	3,144 (21.01)
	2	3,781 (25.27)
	3	2,433 (16.26)
	4	2,102 (14.05)
	5	1,917 (12.81)
	6	775 (5.18)
	≥ 7	809 (5.41)
Annual family income	$0 - $4,999	590 (3.94)
	$5,000 - $9,999	816 (5.45)
	$10,000 - $14,999	933 (6.24)
	$15,000 - $19,999	917 (6.13)
	$20,000 - $24,999	1,106 (7.39)
	$25,000 - $34,999	1,463 (9.78)
	$35,000 - $44,999	1,511 (10.10)
	$45,000 - $54,999	1,055 (7.05)
	$55,000 - $64,999	997 (6.66)
	$65,000 - $74,999	591 (3.95)
	Over $20,000	482 (3.22)
	Under $20,000	128 (0.85)
	$75,000 - $99,999	1,316 (8.80)
	≥ $100,000	2,078 (13.89)
	Refused or don't know	978 (6.54)
Family PIR		2.47 ± 1.62

NHANES: National Health and Nutrition Examination Survey; PIR: poverty income ratio.

healthcare over the past year were important factors [Figure 1]. For the PFOS [Figure 2], PFHxS [Figure 3], PFNA [Figure 4], and ΣPFAS [Figure 5] models, household food security measures were associated with serum concentrations of these PFAS, though the relationships were not clear. For the PFHxS model, telling

Table 2. Legacy serum PFAS distributions (mg/dL), NHANES 1999-2018 (*N* = 14,961)

	Min	Max	P25	P50	P75	IQR
PFOA	0.07	123.00	1.57	2.70	4.40	2.83
PFOS	0.14	1,403.00	5.00	10.20	19.50	14.50
PFHxS	0.07	82.00	0.80	1.50	2.60	1.80
PFNA	0.058	80.77	0.50	0.83	1.39	0.89
ΣPFAS	0.34	1,418.30	9.00	16.40	28.30	19.30

PFAS: Per- and polyfluoroalkyl substances; NHANES: National Health and Nutrition Examination Survey; Min: minimum; Max: maximum; P25: first quartile; P50: median; p75: third quartile; IQR: interquartile range; PFOA: perfluorooctanoic acid; PFOS: perfluorooctanesulfonic acid; PFHxs: perfluorohexane sulfonate; PFNA: perfluorononanoic acid; ΣPFAS: the simple sum of PFOA, PFOS, PFNA, and PFHxS.

a doctor about trouble sleeping was an important factor and was associated with higher serum PFHxS concentrations [Figure 3].

Factors in the home were important in the PFOS [Figure 2] and PFNA [Figure 4] models. Home missing space ownership status was an important factor in the PFOS model, with renting or having another housing arrangement associated with lower serum PFOS concentrations relative to home ownership [Figure 2]. Having a smoker in the home was associated with lower serum concentrations of PFNA [Figure 4].

Discussion

Exogenous factors associated with PFAS concentrations

In this exploratory analysis, we evaluated over 100 variables collected through questionnaires over nine NHANES cycles between 1999 and 2018 as potential factors associated with PFAS concentrations and presented the top 10 factors for four legacy PFAS and ΣPFAS. We observed that year of survey, demographic, income, place of birth and citizenship status, household size, dietary, health, food insecurity, general health and healthcare, housing, social and behavioral, and other characteristics were important factors associated with legacy PFAS concentrations in this nationally representative study population. The most important factors associated with PFAS concentrations varied by congener, suggesting differing potential sources of exposure to each chemical and differences in toxicokinetics affecting body burden.

Year of collection

In general, exposure to certain legacy PFAS, especially PFOS and PFOA, has been declining over time in the U.S. due to phase outs by manufacturers[63,64], and studies have observed declining concentrations of legacy PFAS over time in humans[65,66]. Our findings support this, as the cycle (year) of NHANES was negatively associated with serum PFOA, PFOS, PFHxS, PFNA, and ΣPFAS concentrations. This may be due to diminishing population-level environmental exposure post-ban.

Demographics and family characteristics

Age

The literature surrounding the relationship between age and legacy PFAS concentrations is mixed, though we observed age to be positively associated with serum PFAS concentrations, albeit with relatively small effect sizes. Legacy PFAS bind to serum albumin[67] and albumin concentrations decrease with age[68]. Additionally, glomerular filtration rate, kidney function, and activity level tend to decrease with age[69-71], so a general increase in legacy PFAS concentrations at older ages could be expected[72]. Contraception use, hysterectomies, and menopause may reduce menstrual blood loss as women age and therefore reduce or eliminate a route of PFAS excretion[73], which could increase PFAS concentrations with age, which supports observations in this study. Though parity has been associated with PFAS concentrations in women of reproductive age[23], in elderly women, it has been shown to have a minor effect on PFAS body burden[74].

Table 3. Hypothesis testing summary of the top 10 variables associated with serum PFAS concentrations, by PFAS congener, NHANES 1999-2018

PFAS	Factor	Reference group (categorical variables only)	S/NS	Direction of association (for categorical variables, compared to reference group) (+/-)
PFOA	Year	-	S	-
	Male sex	Female sex	S	+
	Age	-	S	+
	Poverty index	-	S	+
	Born in U.S.	Born outside of U.S.	S	+
	Household size	-	S	-
	Receive healthcare at clinic	Receive healthcare some other place (i.e., not a clinic, doctor's office, ER, or outpatient center)	NS	+
	Receive healthcare at doctor's office		NS	-
	Receive healthcare at ER		NS	-
	Receive healthcare at outpatient center		NS	+
	Emergency food received sometimes	Emergency food received often	S	+
	Food insecurity experienced sometimes	Food insecurity experienced often	NS	+
	Food insecurity experienced never		NS	-
	Health care visits: once in last year	Health care visits: zero in last year	NS	+
	Health care visits: 2-3 times in last year		NS	+
	Health care visits: 4-9 times in last year		NS	-
	Health care visits: 10-12 times in last year		S	-
	Health care visits: > 12 times in last year		S	-
PFOS	Year	-	S	-
	Male sex	Female sex	S	+
	Age	-	S	+
	Household size	-	S	-
	Shellfish consumption in last 30 days	No shellfish consumption in last 30 days	S	+
	Food insecurity experienced sometimes	Food insecurity experienced often	NS	-
	Food insecurity experienced never		NS	-
	Mexican ethnicity	NH White race	NS	-
	Other Hispanic ethnicity		NS	+
	NH Black race		S	+
	NH Asian race		S	+
	Other race		S	+
	Rent home	Other arrangement	NS	+
	Own home		S	+
PFHxS	Year	-	S	-
	Male sex	Female sex	S	+
	Receive retirement income	Do not receive retirement income	S	+
	Born in U.S.	Born outside of U.S.	S	+
	Ever told doctor trouble sleeping	Never told doctor trouble sleeping	S	+

	Married	Living with partner	NS	+
	Widowed		S	+
	Divorced		NS	+
	Separated		NS	-
	Never married		NS	+
	Receive healthcare at clinic	Receive healthcare some other place (i.e., not a clinic, doctor's office, ER, or outpatient center)	NS	-
	Receive healthcare at doctor's office		NS	+
	Receive healthcare at ER		NS	-
	Receive healthcare at outpatient center		NS	-
	Income level $75,000 - $99,000	Income level < $0 - $4,999	S	-
	Income level < $2,000	Income level < $0 - $4,999	S	+
	Receives retirement income	Does not receive retirement income	S	+
PFNA	Year	-	S	-
	Male sex	Female sex	S	+
	Age	-	S	+
	Smoker in home	No smokers in home	NS	+
	Fish consumption in last 30 days	No fish consumption in last 30 days	S	+
	Shellfish consumption in last 30 days	No shellfish consumption in last 30 days	S	+
	Freshwater fish consumed in last 30 days	No freshwater fish consumption in last 30 days	S	+
	Food insecurity experienced sometimes	Food insecurity experienced often	S	-
	Food insecurity experienced never		NS	-
	Fast duration (in hours) before blood draw		S	-
	Mexican ethnicity	NH White race	S	-
	Other Hispanic ethnicity		NS	-
	NH Black race		S	+
	NH Asian race		S	+
	Other race		NS	+
ΣPFAS	Year	-	S	-
	Male sex	Female sex	S	+
	Age	-	S	+
	Born in U.S.	Born outside of U.S.	S	+
	U.S. citizen	Not a U.S. citizen	S	+
	Household size	-	S	-
	Fish consumption in last 30 days	No fish consumption in last 30 days	S	+
	Shellfish consumption in last 30 days	No shellfish consumption in last 30 days	S	+
	Food insecurity experienced sometimes	Food insecurity experienced often	NS	-
	Food insecurity experienced never		S	-
	Mexican ethnicity	NH White race	NS	-
	Other Hispanic ethnicity		NS	+
	NH Black race		S	+
	NH Asian race		S	+
	Other race		NS	+

PFAS: Per- and polyfluoroalkyl substances; NHANES: National Health and Nutrition Examination Survey; S: significant; NS: non-significant; +: positive; -: negative; PFOA: perfluorooctanoic acid; U.S.: United States; ER: emergency room; PFOS: perfluoroctane sulfonate; NH: non-Hispanic; PFHxS: perfluorohexane sulfonate; PFNA: perfluorononanoic acid; ΣPFAS: the simple sum of PFOA, PFOS, PFNA, and PFHxS.

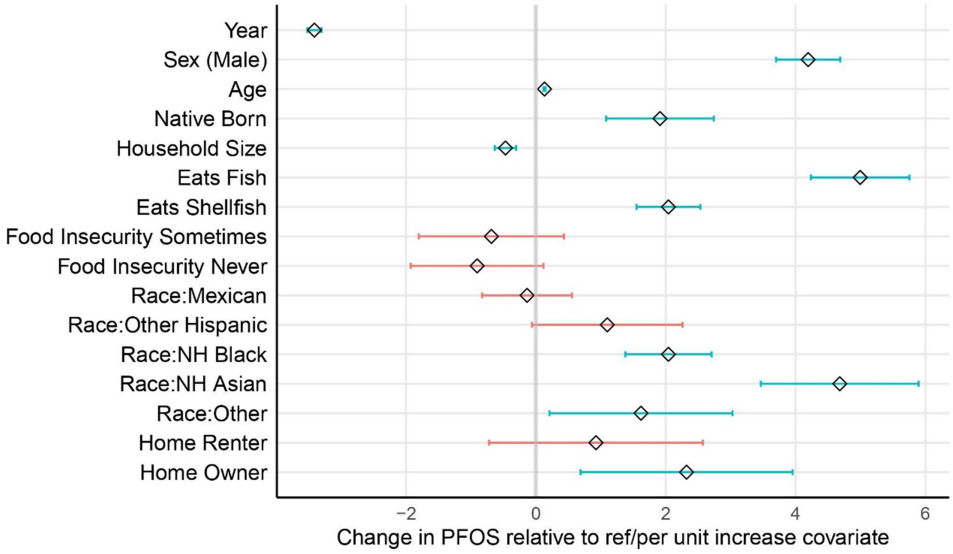

Figure 2. Change in serum PFOS relative to the reference level (for categorical variables) or per unit SD increase for continuous covariates from multivariable mixed-effects regression. See Supplementary Table 2 for details. Blue indicates statistically significant findings, while orange indicates findings that are not statistically significant. PFOS: Perfluorooctane sulfonate.

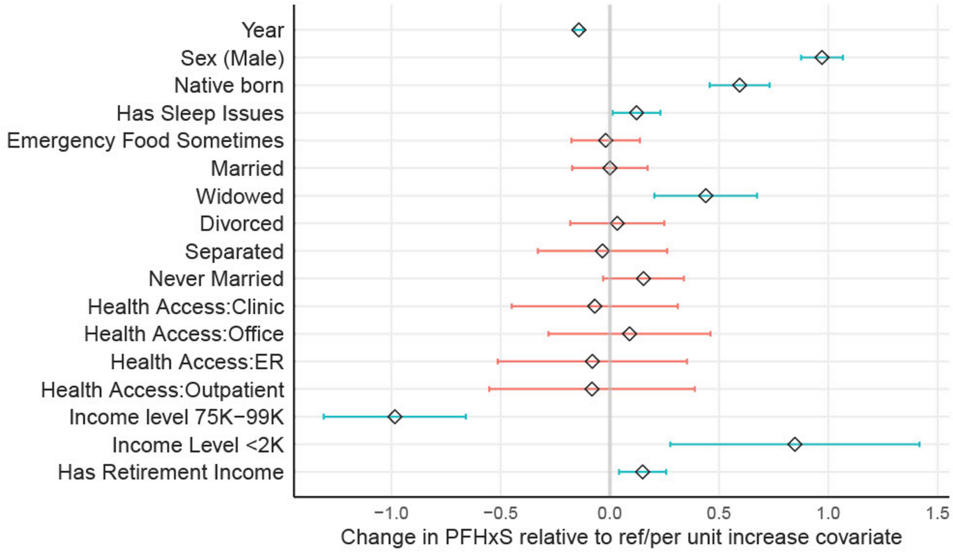

Figure 3. Change in serum PFHxS relative to the reference level (for categorical variables) or per unit SD increase for continuous covariates from multivariable mixed-effects regression. See Supplementary Table 2 for details. Blue indicates statistically significant findings, while orange indicates findings that are not statistically significant. PFHxS: Perfluorohexane sulfonate.

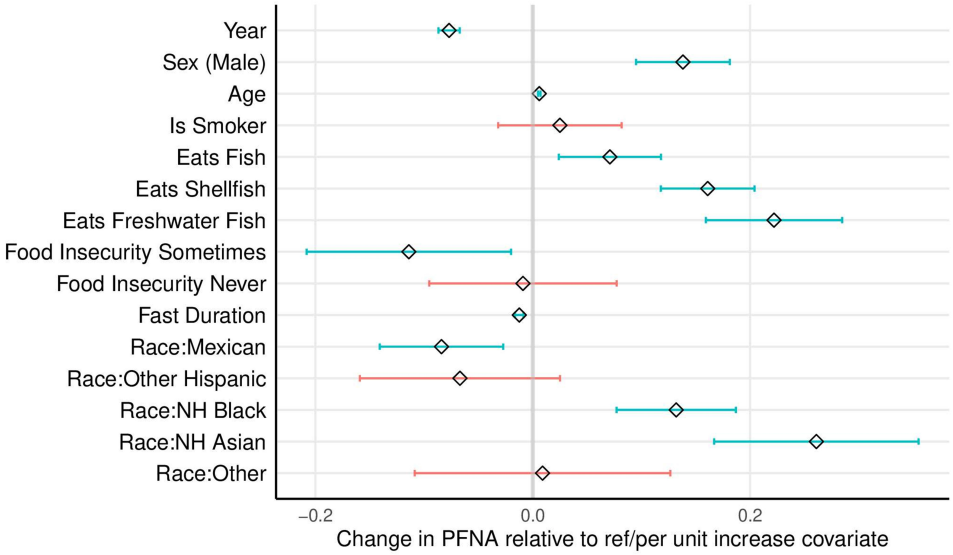

Figure 4. Change in serum PFNA relative to the reference level (for categorical variables) or per unit SD increase for continuous covariates from multivariable mixed-effects regression. See Supplementary Table 2 for details. Blue indicates statistically significant findings, while orange indicates findings that are not statistically significant. PFNA: Perfluorononanoic acid.

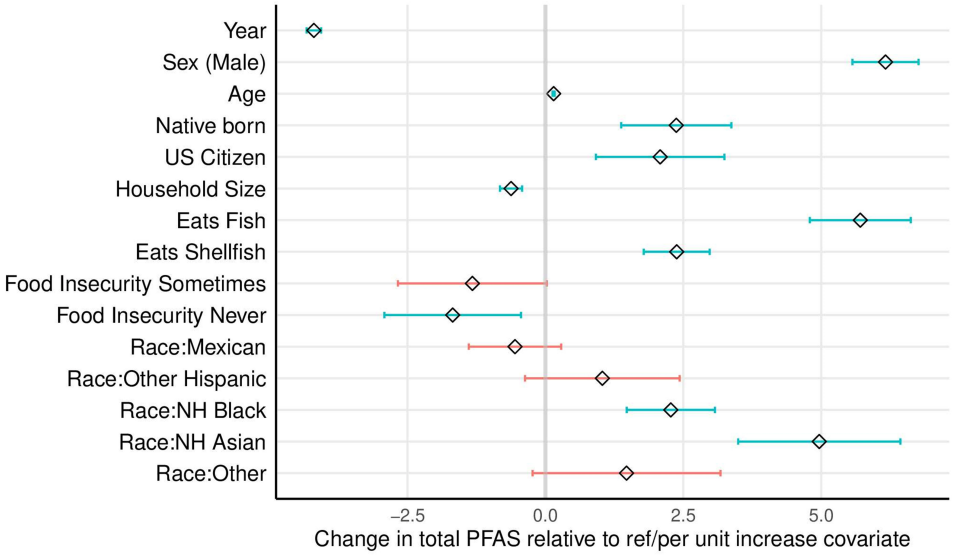

Figure 5. Change in serum ΣPFAS relative to the reference level (for categorical variables) or per unit SD increase for continuous covariates from multivariable mixed-effects regression. See Supplementary Table 2 for details. Blue indicates statistically significant findings, while orange indicates findings that are not statistically significant. ΣPFAS: the simple sum of PFOA, PFOS, PFNA, and PFHxS. PFOA: perfluorooctanoic acid; PFOS: perfluorooctane sulfonate; PFNA: perfluorononanoic acid; PFHxS: perfluorohexane sulfonate.

Citizenship status and place of birth

Though PFOS and PFOA have been phased out in the U.S., they are still in production in parts of the world. In this analysis, we observed U.S. citizenship and being born in the U.S. to be associated with higher concentrations of legacy PFAS concentrations compared to not having U.S. citizenship and being born outside of the U.S. This may be explained in part by cultural habits, imported foods, use of consumer products with PFAS, and social and behavioral differences between immigrants and U.S.-born individuals[75]. Additionally, developed countries may have wider PFAS contamination[76]. A recent review of

determinants of PFAS concentrations in pregnant mothers and neonates found nationality, race, and country of origin to be determinants of legacy PFAS concentrations in pregnant mothers[23].

Household size

Household size was negatively associated with legacy PFAS concentrations in this analysis, which may have several explanations. Mothers giving birth to multiple children would be expected to have lower serum PFAS concentrations, as parity is a factor associated with lower legacy PFAS concentrations[23]. Additionally, this could be a proxy measure for socioeconomic status, as larger family size has been a positive predictor of poverty[77] or household members may not be family but roommates. Given our mixed results related to socioeconomic status and serum PFAS concentrations, further study is needed to better understand this finding.

Sex

Overall, male sex predicted higher concentrations of all evaluated PFAS relative to females. This supports a large body of evidence suggesting females have lower serum PFAS concentrations than males, particularly during reproductive years. Females have greater routes of excretion of PFAS, including menstruation, breastfeeding, and childbirth[78-81]. Additionally, a higher renal clearance of PFAS in females has been suggested, and longer half-lives have been observed in males[82,83].

Race/ethnicity

Race/ethnicity were associated with serum concentrations of PFOA, PFOS, and PFNA concentrations. Compared to the NH White race/ethnicity group, the Mexican-American race/ethnicity group generally had lower PFAS concentrations, while other Hispanic, NH Black, "Other" race, and NH Asian race/ethnicity groups had elevated serum concentrations of these congeners. NH Asian race was associated with the highest serum concentrations. Some studies have documented racial disparities in PFAS exposure, with NH Black individuals having higher concentrations of PFOS; however, studies in recent years have shown changes in these trends[45,84-86]. A recent study observed community water systems serving higher proportions of Hispanic/Latino and NH Black residents had significantly increased odds of detecting several PFAS[87], which may reflect disparities in the sites of point sources of PFAS contamination. Our findings in this study suggest a complex relationship between socioeconomic status and serum PFAS concentrations.

Differences in tap water consumption may partially explain our findings. A study assessing NHANES 2011-2016 water consumption trends observed NH White individuals consumed the most tap water, whereas Mexican-American individuals consumed the most bottled water[88], with other studies observing similar results[89-91]. Although bottled water can contain detectable quantities of PFAS[92-94], there are many cases of contaminated public drinking water supplies in the U.S.[95]. The higher concentrations of PFOA, PFOS, and PFNA predicted by NH Asian race compared to other racial and ethnic groups contradict some findings in other studies[21].

Diet

Dietary freshwater fish

Consumption of freshwater fish, such as walleye, trout, pike, perch, or bass, in the last 30 days was significantly associated with higher concentrations of PFOS, PFNA, and ΣPFAS. This is supported by findings from other studies, which have found freshwater fish consumption to be a significant contributor to PFOS concentrations[96-99]. It has been suggested that even infrequent freshwater fish consumption may increase serum PFOS concentrations[97]. Though our findings are based on a binary measure of consumption (i.e., ever *vs.* never in the last 30 days), this suggests that even at a potentially wide range of

freshwater fish consumption levels, there is exposure to legacy PFAS. The risks and benefits of freshwater fish consumption may warrant further study.

Dietary shellfish and fish

Shellfish and fish consumption was associated with higher serum concentrations of PFOS, PFNA and ΣPFAS [Figures 2, 4, and 5]. This supports other literature finding evidence of an association between shellfish consumption and serum PFAS concentrations[100-105]. Because our findings are based on a binary measure of consumption, further research in populations consuming shellfish may be helpful in identifying specific species that may result in higher exposure. While NHANES collects information on specific types of shellfish and fish consumed, counts were too small for any specific fish or shellfish to evaluate in this analysis.

Food insecurity

Certain measures of food insecurity (i.e., household food not lasting and household receiving emergency food) were factors associated with all PFAS concentrations, though these relationships were inconsistent. This again suggests a complex relationship between socioeconomic status and serum PFAS concentrations.

Health

Receipt of healthcare

The number of healthcare visits in the last year was associated with PFOA concentrations, though the relationship was not clear. Compared to no visits in the last year, one or two to three visits were associated with higher serum concentrations. Compared to no visits in the last year, 4-9, 10-12, and > 12 visits in the last year were associated with lower concentrations of PFOA. These findings may suggest relationships between overall health, insurance, or socioeconomic status and serum PFAS concentrations. Responding no (*vs.* yes) to "Have you ever told a doctor you had trouble sleeping?" was associated with higher serum concentrations of PFHxS, which may be reflective of healthcare access rather than a relationship between sleeping troubles and serum PFAS concentrations. Additionally, the type of place where healthcare was usually received was associated with PFOA and PFHxS concentrations. Relationships were not clear and may be influenced by healthcare access and socioeconomic status.

Housing and income

Housing characteristics

Renting rather than owning a home was associated with lower concentrations of PFOS. Though associations between housing characteristics and PFAS exposure are not well studied, carpet treatments contain PFAS and have been associated with elevated PFAS concentrations[106,107]. House dust has also been suggested to be an important pathway to human PFAS exposure[108].

Income and consumer spending

Family poverty income ratio (PIR) was positively associated with serum PFOA concentrations, receiving retirement income was positively associated with PFHxS concentrations, and while there was an association between household income and PFHxS concentrations, the relationship did not appear to be linear. The first two measures suggest a positive relationship between income and PFAS concentrations, which may be explained by the increased opportunity to use legacy PFAS containing products such as clothing at higher prices[109].

Comparison to other studies

The R^2 values in this analysis were relatively low, with the highest at 0.31; however, our R^2 results are comparable to or higher than other studies evaluating factors associated with PFAS concentrations[21,26,30,110,111].

Imputation methods

Given the high percent missingness in certain variables and the exploratory nature of this analysis, we tested two donor-based single imputation methods. Imputation involves assumptions that can influence results; therefore, these findings should be interpreted with caution and should be confirmed in future studies. Though the best single imputation methods may vary by dataset and application, research suggests that hot-deck imputation is superior compared to PMM when comparing root mean squared error, unsupervised classification error, supervised classification error, and the time used to run the algorithm[112].

It should be noted that the data in this exploratory analysis had high missingness and a complete-case analysis could not be conducted for comparison due to this and the large number of variables included in models. Further study is needed to understand the best imputation methods for data and modeling of this nature.

Individual congeners vs. additive sums

The recent National Academies of Sciences, Engineering, and Medicine (NASEM) guidance on patient follow-up after PFAS testing was based on ranges of values of the simple additive sum of methylperfluorooctanesulfonamidoacetic acid (MeFOSAA), PFHxS, PFOA, PFDA, perfluoroundecanoic acid (PFUnDA), PFOS, and PFNA in serum or plasma[47]. For patients with < 2 ng/mL serum or plasma PFAS, NASEM recommends providing the usual standard of care. For patients with ≥ 20 ng/mL serum or plasma PFAS, NASEM encourages exposure reduction if point sources can be identified, prioritizing screening for dyslipidemia, conducting thyroid function testing, assessing signs and symptoms of kidney cancer, and assessing signs and symptoms of testicular cancer and ulcerative colitis. Though our analysis could not include MeFOSAA, PFDA, and PFUnDA due to the limited number of NHANES cycles that have evaluated these PFAS, we do show differences in the top factors associated with ΣPFAS (PFOA + PFOS + PFHxS + PFNA) and PFOA, PFOS, PFHxS, and PFNA individually. Therefore, it is important to consider factors associated with simple additive sums of commonly detected PFAS in addition to factors related to single PFAS congeners.

Strengths and limitations

The present study has notable limitations. The first is the cross-sectional nature of the NHANES data. Although legacy PFAS may have half-lives of several years or more[113], which we expect to reduce the risk of exposure misclassification, this does not completely mitigate risk. Additionally, we cannot establish causality and do not draw conclusions as to whether the identified factors cause changes in PFAS concentrations. Second, most of the factors evaluated were collected through self-report. Third, many covariates could not be assessed for their association with PFAS concentrations in this study population due to high missingness. Given the large number of factors that could not be included in this analysis, further research is needed to continue establishing factors associated with PFAS concentrations. Covariates surrounding consumer behavior, diet, housing characteristics, physical activity, and reproductive health, in particular, require further study and may offer further insight into exposure and elimination of PFAS. Fourth, we evaluated the four most measured PFAS, though there are thousands of PFAS. Given the increased use of short-chain and alternative PFAS in recent years, more research is needed to comprehensively understand the factors associated with PFAS concentrations. Fifth, this is an unweighted exploratory study. Although NHANES provides weights for each cycle due to the complex survey design, we did not incorporate them due to the high missingness and methods employed in this exploratory analysis.

Despite these limitations, this work has several notable strengths. First, this was a large study population that was representative of the U.S. general population. Second, this analysis evaluates the largest number of potential factors associated with PFAS concentrations to date. Most studies evaluating factors associated with PFAS concentrations consider a limited number of factors. Third, we used the well-known ENR approach to identify the top factors associated with PFAS concentrations in this study population.

Future directions

It is important to consider the findings here exploratory, and replication in other populations representative of the U.S. general population is needed to confirm these findings. Many of the variables included in this analysis were related to one another or were consequences of others (e.g., poverty index or income and food insecurity). We did not perform stratified analyses due to the exploratory nature of this analysis; therefore, future studies stratified by factors identified here (e.g., socioeconomic status, race, sex, age) are warranted.

Future studies may consider evaluating socioeconomic status in multiple ways (i.e., food insecurity, healthcare access, *etc.*) to further tease out the associations observed here. It may also be beneficial to consider the joint effects of socioeconomic status and PFAS on health effects or socioeconomic status as a moderator in PFAS-health outcome relationships. Given our findings on factors associated with PFAS (i.e., many indicators of socioeconomic status), the potential health effects associated with both PFAS exposure and socioeconomic status, and that PFAS exposure through contaminated drinking water has been associated with socioeconomic and sociodemographic factors[87] and is an important environmental justice issue within the U.S., it is important to explore these relationships further.

This study was motivated by increasing concerns over health effects following PFAS exposure in adults, and our findings raise awareness of potentially vulnerable groups in the general population not occupationally exposed or exposed through highly contaminated drinking water. Future studies should consider prospective analyses where possible to assess potential causality between these factors and serum PFAS concentrations and to further investigate how these factors influence PFAS elimination half-lives. Research should consider adjusting for factors with substantial evidence to suggest they may affect PFAS concentrations, such as age, sex, and year of sample collection, as well as factors specific to the population being studied (i.e., children, adults, pregnant mothers, highly exposed communities, *etc.*). It may be helpful to stratify by factors such as age, sex, year of sample collection, socioeconomic status, and other factors identified here to identify potential effect modifiers in PFAS studies dependent on the research question. Additionally, studies investigating factors associated with PFAS concentrations should consider the inclusion of short-chain and newer PFAS, where little research has been conducted. The R^2 values found in our models suggest other factors may be associated with higher serum PFAS concentrations. Biological factors associated with PFAS concentrations should be explored further, as these factors may affect PFAS toxicokinetics.

CONCLUSIONS

This novel, exploratory study aimed to address the substantial gaps in our understanding of factors associated with serum concentrations of legacy PFAS chemicals in the general U.S. population by employing elastic net regression to evaluate over 100 variables collected via questionnaire in NHANES cycles between 1999-2018. While specific factors varied by the specific PFAS evaluated, we observed the time period of sample collection, sex, age, race, socioeconomic factors, and fish or shellfish consumption to be important factors associated with legacy PFAS concentrations in a large U.S. study population. These preliminary findings may direct research priorities related to PFAS exposure and disparities and highlight gaps in risk assessment. Reducing PFAS exposure through the identification of vulnerable groups may

reduce disease burden in the general population and reduce healthcare costs.

DECLARATIONS

Acknowledgments

The authors would like to thank Sana Savadatti, Elizabeth Lewis-Michl, Steven Forand, and Steve Shost for their review of this work and feedback. We thank Megan Kern of Bassett Research Institute for her help with creating our graphic abstract and pipeline flowchart.

Authors' contributions

Conceptualization: McAdam J, Jones LE, Bell EM, Romeiko XX
Methodology, software, visualization, and formal analysis: McAdam J, Jones LE
Writing - original draft: McAdam J, Jones LE, Bell EM
Writing - review and editing: McAdam J, Jones LE, Bell EM, Romeiko XX
Supervision: Bell EM
All authors discussed the results and contributed to the final manuscript.

Availability of data and materials

NHANES datasets are publicly available: https://www.cdc.gov/nchs/nhanes/index.htm.

Financial support and sponsorship

This work was supported by Centers for Disease Control and Prevention Grant # U01TS000303.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

NHANES data were collected under study protocols approved by the National Center for Health Statistics institutional review board and all participants gave written informed consent. More information is available here: https://www.cdc.gov/nchs/nhanes/irba98.htm.

Consent for publication

Not applicable.

REFERENCES

1. ATSDR. PFAS exposure assessments. Available from: https://www.atsdr.cdc.gov/pfas/activities/assessments.html. [Last accessed on 29 May 2024].
2. CDC. Fourth national report on human exposure to environmental chemicals, updated tables, March 2021: volume three: analysis of pooled serum samples for select chemicals, NHANES 2005-2016. Available from: https://stacks.cdc.gov/view/cdc/105344. [Last accessed on 29 May 2024].
3. Dunder L, Lind PM, Salihovic S, Stubleski J, Kärrman A, Lind L. Changes in plasma levels of per- and polyfluoroalkyl substances (PFAS) are associated with changes in plasma lipids - a longitudinal study over 10 years. *Environ Res* 2022;211:112903. DOI PubMed
4. Li Y, Barregard L, Xu Y, et al. Associations between perfluoroalkyl substances and serum lipids in a Swedish adult population with contaminated drinking water. *Environ Health* 2020;19:33. DOI PubMed PMC
5. Lin PD, Cardenas A, Hauser R, et al. Per- and polyfluoroalkyl substances and blood lipid levels in pre-diabetic adults-longitudinal analysis of the diabetes prevention program outcomes study. *Environ Int* 2019;129:343-53. DOI PubMed PMC
6. Liu JJ, Cui XX, Tan YW, et al. Per- and perfluoroalkyl substances alternatives, mixtures and liver function in adults: a community-based population study in China. *Environ Int* 2022;163:107179. DOI PubMed

7. Salihovic S, Stubleski J, Kärrman A, et al. Changes in markers of liver function in relation to changes in perfluoroalkyl substances - a longitudinal study. *Environ Int* 2018;117:196-203. DOI PubMed
8. Averina M, Brox J, Huber S, Furberg AS. Exposure to perfluoroalkyl substances (PFAS) and dyslipidemia, hypertension and obesity in adolescents. the fit futures study. *Environ Res* 2021;195:110740. DOI PubMed
9. Pitter G, Zare Jeddi M, Barbieri G, et al. Perfluoroalkyl substances are associated with elevated blood pressure and hypertension in highly exposed young adults. *Environ Health* 2020;19:102. DOI PubMed PMC
10. Bartell SM, Vieira VM. Critical review on PFOA, kidney cancer, and testicular cancer. *J Air Waste Manag Assoc* 2021;71:663-79. DOI PubMed
11. Papadopoulou E, Sabaredzovic A, Namork E, Nygaard UC, Granum B, Haug LS. Exposure of Norwegian toddlers to perfluoroalkyl substances (PFAS): the association with breastfeeding and maternal PFAS concentrations. *Environ Int* 2016;94:687-94. DOI PubMed
12. EFSA Panel on Contaminants in the Food Chain (EFSA CONTAM Panel); Schrenk D, Bignami M, Bodin L, et al. Risk to human health related to the presence of perfluoroalkyl substances in food. *EFSA J* 2020;18:e06223. DOI
13. Schaider LA, Balan SA, Blum A, et al. Fluorinated compounds in U.S. fast food packaging. *Environ Sci Technol Lett* 2017;4:105-11. DOI PubMed PMC
14. Glenn G, Shogren R, Jin X, Orts W, Hart-Cooper W, Olson L. Per- and polyfluoroalkyl substances and their alternatives in paper food packaging. *Compr Rev Food Sci Food Saf* 2021;20:2596-625. DOI PubMed
15. Curtzwiler GW, Silva P, Hall A, Ivey A, Vorst K. Significance of perfluoroalkyl substances (PFAS) in food packaging. *Integr Environ Assess Manag* 2021;17:7-12. DOI PubMed
16. Seltenrich N. PFAS in food packaging: a hot, greasy exposure. *Environ Health Perspect* 2020;128:54002. DOI PubMed PMC
17. ATSDR. Toxicological profile for perfluoroalkyls. 2021. Available from: https://www.atsdr.cdc.gov/toxprofiles/tp200.pdf. [Last accessed on 29 May 2024].
18. Thépaut E, Dirven HAAM, Haug LS, et al. Per- and polyfluoroalkyl substances in serum and associations with food consumption and use of personal care products in the Norwegian biomonitoring study from the EU project EuroMix. *Environ Res* 2021;195:110795. DOI
19. Trudel D, Horowitz L, Wormuth M, Scheringer M, Cousins IT, Hungerbühler K. Estimating consumer exposure to PFOS and PFOA. *Risk Anal* 2008;28:251-69. DOI PubMed
20. Kato K, Wong LY, Chen A, et al. Changes in serum concentrations of maternal poly- and perfluoroalkyl substances over the course of pregnancy and predictors of exposure in a multiethnic cohort of Cincinnati, Ohio pregnant women during 2003-2006. *Environ Sci Technol* 2014;48:9600-8. DOI PubMed PMC
21. Park SK, Peng Q, Ding N, Mukherjee B, Harlow SD. Determinants of per- and polyfluoroalkyl substances (PFAS) in midlife women: evidence of racial/ethnic and geographic differences in PFAS exposure. *Environ Res* 2019;175:186-99. DOI PubMed PMC
22. Sagiv SK, Rifas-Shiman SL, Webster TF, et al. Sociodemographic and perinatal predictors of early pregnancy per- and polyfluoroalkyl substance (PFAS) concentrations. *Environ Sci Technol* 2015;49:11849-58. DOI PubMed PMC
23. McAdam J, Bell EM. Determinants of maternal and neonatal PFAS concentrations: a review. *Environ Health* 2023;22:41. DOI PubMed PMC
24. Nyström J, Benskin JP, Plassmann M, et al. Demographic, life-style and physiological determinants of serum per- and polyfluoroalkyl substance (PFAS) concentrations in a national cross-sectional survey of Swedish adolescents. *Environ Res* 2022;208:112674. DOI PubMed
25. Bartolomé M, Gallego-Picó A, Cutanda F, et al; Bioambient.es. Perfluorinated alkyl substances in Spanish adults: geographical distribution and determinants of exposure. *Sci Total Environ* 2017;603-4:352-60. DOI PubMed
26. Barton KE, Starling AP, Higgins CP, McDonough CA, Calafat AM, Adgate JL. Sociodemographic and behavioral determinants of serum concentrations of per- and polyfluoroalkyl substances in a community highly exposed to aqueous film-forming foam contaminants in drinking water. *Int J Hyg Environ Health* 2020;223:256-66. DOI PubMed PMC
27. Richterová D, Govarts E, Fábelová L, et al. PFAS levels and determinants of variability in exposure in European teenagers - Results from the HBM4EU aligned studies (2014-2021). *Int J Hyg Environ Health* 2023;247:114057. DOI PubMed PMC
28. Pitter G, Da Re F, Canova C, et al. Serum levels of perfluoroalkyl substances (PFAS) in adolescents and young adults exposed to contaminated drinking water in the Veneto Region, Italy: a cross-sectional study based on a health surveillance program. *Environ Health Perspect* 2020;128:27007. DOI PubMed PMC
29. Li Y, Andersson A, Xu Y, et al. Determinants of serum half-lives for linear and branched perfluoroalkyl substances after long-term high exposure - a study in Ronneby, Sweden. *Environ Int* 2022;163:107198. DOI PubMed
30. Cuffney M, Wilkie AA, Kotlarz N, et al. Factors associated with per- and polyfluoroalkyl substances (PFAS) serum concentrations in residents of New Hanover County, North Carolina: the GenX exposure study. *Environ Res* 2023;237:117020. DOI PubMed PMC
31. Fábelová L, Beneito A, Casas M, et al. PFAS levels and exposure determinants in sensitive population groups. *Chemosphere* 2023;313:137530. DOI PubMed PMC
32. Tsai MS, Miyashita C, Araki A, et al. Determinants and temporal trends of perfluoroalkyl substances in pregnant women: the Hokkaido study on environment and children's health. *Int J Environ Res Public Health* 2018;15:989. DOI PubMed PMC
33. Tian Y, Zhou Y, Miao M, et al. Determinants of plasma concentrations of perfluoroalkyl and polyfluoroalkyl substances in pregnant women from a birth cohort in Shanghai, China. *Environ Int* 2018;119:165-73. DOI PubMed

34. Huo X, Liang W, Tang W, et al; Shanghai Birth Cohort Study. Dietary and maternal sociodemographic determinants of perfluoroalkyl and polyfluoroalkyl substance levels in pregnant women. *Chemosphere* 2023;332:138863. DOI PubMed
35. Mahfouz M, Harmouche-Karaki M, Matta J, et al. Maternal serum, cord and human milk levels of per- and polyfluoroalkyl substances (PFAS), association with predictors and effect on newborn anthropometry. *Toxics* 2023;11:455. DOI PubMed PMC
36. Sørensen MM, Fisker AB, Dalgård C, et al. Predictors of serum- per- and polyfluoroalkyl substance (PFAS) concentrations among infants in Guinea-Bissau, West Africa. *Environ Res* 2023;228:115784. DOI PubMed
37. Manzano-Salgado CB, Casas M, Lopez-Espinosa MJ, et al. Variability of perfluoroalkyl substance concentrations in pregnant women by socio-demographic and dietary factors in a Spanish birth cohort. *Environ Int* 2016;92-3:357-65. DOI PubMed
38. Jain RB, Ducatman A. Factors affecting serum PFAS concentrations among US females with surgically and naturally induced menopause: data from NHANES 2003-2018. *Environ Sci Pollut Res Int* 2023;30:84705-24. DOI PubMed
39. Obeng-Gyasi E. Factors associated with elevated per- and polyfluoroalkyl substances serum levels in older adults. *Aging Health Res* 2022;2:100086. DOI PubMed PMC
40. Graber JM, Black TM, Shah NN, et al. Prevalence and predictors of per- and polyfluoroalkyl substances (PFAS) serum levels among members of a suburban US volunteer fire department. *Int J Environ Res Public Health* 2021;18:3730. DOI PubMed PMC
41. Chiu WA, Lynch MT, Lay CR, et al. Bayesian estimation of human population toxicokinetics of PFOA, PFOS, PFHxS, and PFNA from studies of contaminated drinking water. *Environ Health Perspect* 2022;130:127001. DOI PubMed PMC
42. Seshasayee SM, Rifas-Shiman SL, Chavarro JE, et al. Dietary patterns and PFAS plasma concentrations in childhood: project Viva, USA. *Environ Int* 2021;151:106415. DOI PubMed PMC
43. CDC. National health and nutrition examination survey: analytic guidelines, 1999-2010. Available from: https://www.cdc.gov/nchs/data/series/sr_02/sr02_161.pdf. [Last accessed on 29 May 2024].
44. Calafat AM, Wong LY, Kuklenyik Z, Reidy JA, Needham LL. 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;115:1596-602. DOI PubMed PMC
45. Kato K, Wong LY, Jia LT, Kuklenyik Z, Calafat AM. Trends in exposure to polyfluoroalkyl chemicals in the U.S. population: 1999-2008. *Environ Sci Technol* 2011;45:8037-45. DOI PubMed
46. Hornung RW, Reed LD. Estimation of average concentration in the presence of nondetectable values. *Appl Occup Environ Hyg* 1990;5:46-51. DOI
47. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Division on Earth and Life Studies; Board on Population Health and Public Health Practice; Board on Environmental Studies and Toxicology; Committee on the Guidance on PFAS Testing and Health Outcomes. Guidance on PFAS exposure, testing, and clinical follow-up. 2022. DOI PubMed
48. Analytic and reporting guidelines: the national health and nutrition examination survey (NHANES). 2005. Available from: https://wwwn.cdc.gov/nchs/data/nhanes/analyticguidelines/05-06-analytic-guidelines.pdf. [Last accessed on 29 May 2024].
49. Zhao Y, Jin H, Qu J, et al. The influences of perfluoroalkyl substances on the rheumatoid arthritis clinic. *BMC Immunol* 2022;23:10. DOI PubMed PMC
50. von Holst H, Nayak P, Dembek Z, et al. Perfluoroalkyl substances exposure and immunity, allergic response, infection, and asthma in children: review of epidemiologic studies. *Heliyon* 2021;7:e08160. DOI PubMed PMC
51. Stanifer JW, Stapleton HM, Souma T, Wittmer A, Zhao X, Boulware LE. Perfluorinated chemicals as emerging environmental threats to kidney health: a scoping review. *Clin J Am Soc Nephrol* 2018;13:1479-92. DOI PubMed PMC
52. Meneguzzi A, Fava C, Castelli M, Minuz P. Exposure to perfluoroalkyl chemicals and cardiovascular disease: experimental and epidemiological evidence. *Front Endocrinol* 2021;12:706352. DOI PubMed PMC
53. Kim MJ, Moon S, Oh BC, et al. Association between perfluoroalkyl substances exposure and thyroid function in adults: a meta-analysis. *PLoS One* 2018;13:e0197244. DOI PubMed PMC
54. Gui SY, Qiao JC, Xu KX, et al. Association between per- and polyfluoroalkyl substances exposure and risk of diabetes: a systematic review and meta-analysis. *J Expo Sci Environ Epidemiol* 2023;33:40-55. DOI PubMed
55. Fenton SE, Ducatman A, Boobis A, et al. Per- and polyfluoroalkyl substance toxicity and human health review: current state of knowledge and strategies for informing future research. *Environ Toxicol Chem* 2021;40:606-30. DOI PubMed PMC
56. Kuhn M, Johnson K. Handling missing data. In: Feature engineering and selection: a practical approach for predictive models. Chapman and Hall; 2019. DOI
57. Emmanuel T, Maupong T, Mpoeleng D, Semong T, Mphago B, Tabona O. A survey on missing data in machine learning. *J Big Data* 2021;8:140. DOI PubMed PMC
58. HotDeckImputation-package: hot deck imputation methods for missing data. Available from: https://www.rdocumentation.org/packages/HotDeckImputation/versions/1.1.0/topics/HotDeckImputation-package. [Last accessed on 29 May 2024].
59. Andridge RR, Little RJ. A review of hot deck imputation for survey non-response. *Int Stat Rev* 2010;78:40-64. DOI PubMed PMC
60. Kowarik A, Templ M. Imputation with the R Package VIM. *J Stat Soft* 2016;74:1-16. DOI
61. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Soft* 2011;45:1-67. DOI
62. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J R Stat Soc B* 2005;67:301-20. DOI
63. USEPA. EPA and 3M announce phase out of pfos. 2000. Available from: https://www.epa.gov/archive/epapages/newsroom_archive/newsreleases/33aa946e6cb11f35852568e1005246b4.html. [Last accessed on 29 May 2024].
64. USEPA. PFOA stewardship program baseline year summary report. Available from: https://www.epa.gov/assessing-and-managing-

chemicals-under-tsca/pfoa-stewardship-program-baseline-year-summary-report. [Last accessed on 29 May 2024].
65. Toms LM, Thompson J, Rotander A, et al. Decline in perfluorooctane sulfonate and perfluorooctanoate serum concentrations in an Australian population from 2002 to 2011. *Environ Int* 2014;71:74-80. DOI PubMed PMC
66. Hurley S, Goldberg D, Wang M, et al. Time trends in per- and polyfluoroalkyl substances (PFASs) in California women: declining serum levels, 2011-2015. *Environ Sci Technol* 2018;52:277-87. DOI PubMed
67. Salvalaglio M, Muscionico I, Cavallotti C. Determination of energies and sites of binding of PFOA and PFOS to human serum albumin. *J Phys Chem B* 2010;114:14860-74. DOI PubMed
68. Weaving G, Batstone GF, Jones RG. Age and sex variation in serum albumin concentration: an observational study. *Ann Clin Biochem* 2016;53:106-11. DOI PubMed
69. Denic A, Glassock RJ, Rule AD. Structural and functional changes with the aging kidney. *Adv Chronic Kidney Dis* 2016;23:19-28. DOI PubMed PMC
70. Milanović Z, Pantelić S, Trajković N, Sporiš G, Kostić R, James N. Age-related decrease in physical activity and functional fitness among elderly men and women. *Clin Interv Aging* 2013;8:549-56. DOI PubMed PMC
71. Musso CG, Oreopoulos DG. Aging and physiological changes of the kidneys including changes in glomerular filtration rate. *Nephron Physiol* 2011;119 Suppl 1:p1-5. DOI PubMed
72. Verner MA, Loccisano AE, Morken NH, et al. Associations of perfluoroalkyl substances (PFAS) with lower birth weight: an evaluation of potential confounding by glomerular filtration rate using a physiologically based pharmacokinetic model (PBPK). *Environ Health Perspect* 2015;123:1317-24. DOI PubMed PMC
73. Rickard BP, Rizvi I, Fenton SE. Per- and poly-fluoroalkyl substances (PFAS) and female reproductive outcomes: PFAS elimination, endocrine-mediated effects, and disease. *Toxicology* 2022;465:153031. DOI PubMed PMC
74. Salihovic S, Kärrman A, Lind L, Lind PM, Lindström G, van Bavel B. Perfluoroalkyl substances (PFAS) including structural PFOS isomers in plasma from elderly men and women from Sweden: results from the prospective investigation of the vasculature in Uppsala seniors (PIVUS). *Environ Int* 2015;82:21-7. DOI PubMed
75. Muennig P, Song X, Payne-Sturges DC, Gee GC. Blood and urine levels of long half-life toxicants by nativity among immigrants to the United States. *Sci Total Environ* 2011;412-3:109-13. DOI PubMed
76. Fiedler H, Sadia M. Regional occurrence of perfluoroalkane substances in human milk for the global monitoring plan under the Stockholm Convention on Persistent Organic Pollutants during 2016-2019. *Chemosphere* 2021;277:130287. DOI PubMed
77. Rahman MA. Household characteristics and poverty: a logistic regression analysis. *J Develop Areas* 2013;47:303-17. Available from: http://www.jstor.org/stable/23612271. [Last accessed on 29 May 2024]
78. Cariou R, Veyrand B, Yamada A, et al. Perfluoroalkyl acid (PFAA) levels and profiles in breast milk, maternal and cord serum of French women and their newborns. *Environ Int* 2015;84:71-81. DOI PubMed
79. Lorber M, Eaglesham GE, Hobson P, Toms LM, Mueller JF, Thompson JS. The effect of ongoing blood loss on human serum concentrations of perfluorinated acids. *Chemosphere* 2015;118:170-7. DOI PubMed
80. Manzano-Salgado CB, Casas M, Lopez-Espinosa MJ, et al. Transfer of perfluoroalkyl substances from mother to fetus in a Spanish birth cohort. *Environ Res* 2015;142:471-8. DOI PubMed
81. Mondal D, Weldon RH, Armstrong BG, et al. Breastfeeding: a potential excretion route for mothers and implications for infant exposure to perfluoroalkyl acids. *Environ Health Perspect* 2014;122:187-92. DOI PubMed PMC
82. Fromme H, Tittlemier SA, Völkel W, Wilhelm M, Twardella D. Perfluorinated compounds - exposure assessment for the general population in Western countries. *Int J Hyg Environ Health* 2009;212:239-70. DOI PubMed
83. Li Y, Fletcher T, Mucs D, et al. Half-lives of PFOS, PFHxS and PFOA after end of exposure to contaminated drinking water. *Occup Environ Med* 2018;75:46-51. DOI PubMed PMC
84. CDC. Fourth national report on human exposure to environmental chemicals: updated tables, January 2019, volume one. 2019. DOI
85. Jain RB. Contribution of diet and other factors to the levels of selected polyfluorinated compounds: data from NHANES 2003-2008. *Int J Hyg Environ Health* 2014;217:52-61. DOI PubMed
86. Nelson JW, Scammell MK, Hatch EE, Webster TF. Social disparities in exposures to bisphenol A and polyfluoroalkyl chemicals: a cross-sectional study within NHANES 2003-2006. *Environ Health* 2012;11:10. DOI PubMed PMC
87. Liddie JM, Schaider LA, Sunderland EM. Sociodemographic factors are associated with the abundance of PFAS sources and detection in U.S. community water systems. *Environ Sci Technol* 2023;57:7902-12. DOI PubMed PMC
88. Vieux F, Maillot M, Rehm CD, Barrios P, Drewnowski A. Trends in tap and bottled water consumption among children and adults in the United States: analyses of NHANES 2011-16 data. *Nutr J* 2020;19:10. DOI PubMed PMC
89. Gorelick MH, Gould L, Nimmer M, et al. Perceptions about water and increased use of bottled water in minority children. *Arch Pediatr Adolesc Med* 2011;165:928-32. DOI PubMed
90. Rosinger AY, Herrick KA, Wutich AY, Yoder JS, Ogden CL. Disparities in plain, tap and bottled water consumption among US adults: national health and nutrition examination survey (NHANES) 2007-2014. *Public Health Nutr* 2018;21:1455-64. DOI PubMed PMC
91. Park S, Onufrak S, Patel A, Sharkey JR, Blanck HM. Perceptions of drinking water safety and their associations with plain water intake among US Hispanic adults. *J Water Health* 2019;17:587-96. DOI PubMed PMC
92. Consumer Reports. What's really in your bottled water? 2020. Available from: https://www.consumerreports.org/water-quality/whats-really-in-your-bottled-water-a5361150329/. [Last accessed on 29 May 2024].

93. Chow SJ, Ojeda N, Jacangelo JG, Schwab KJ. Detection of ultrashort-chain and other per- and polyfluoroalkyl substances (PFAS) in U.S. bottled water. *Water Res* 2021;201:117292. DOI PubMed
94. Krachler M, Shotyk W. Trace and ultratrace metals in bottled waters: survey of sources worldwide and comparison with refillable metal bottles. *Sci Total Environ* 2009;407:1089-96. DOI PubMed
95. Stoiber T, Evans S, Temkin AM, Andrews DQ, Naidenko OV. PFAS in drinking water: an emergent water quality threat. 2020. Available from: https://www.ewg.org/sites/default/files/u352/Stoiber_Evans_WaterSolutions_2020.pdf. [Last accessed on 29 May 2024].
96. Augustsson A, Lennqvist T, Osbeck CMG, et al. Consumption of freshwater fish: a variable but significant risk factor for PFOS exposure. *Environ Res* 2021;192:110284. DOI PubMed
97. Barbo N, Stoiber T, Naidenko OV, Andrews DQ. Locally caught freshwater fish across the United States are likely a significant source of exposure to PFOS and other perfluorinated compounds. *Environ Res* 2023;220:115165. DOI PubMed
98. von Stackelberg K, Li M, Sunderland E. Results of a national survey of high-frequency fish consumers in the United States. *Environ Res* 2017;158:126-36. DOI PubMed
99. Wattigney WA, Savadatti SS, Liu M, et al. Biomonitoring of per- and polyfluoroalkyl substances in minority angler communities in central New York State. *Environ Res* 2022;204:112309. DOI PubMed PMC
100. Christensen KY, Raymond M, Blackowicz M, et al. Perfluoroalkyl substances and fish consumption. *Environ Res* 2017;154:145-51. DOI PubMed
101. Haug LS, Thomsen C, Brantsaeter AL, et al. Diet and particularly seafood are major sources of perfluorinated compounds in humans. *Environ Int* 2010;36:772-8. DOI PubMed
102. Lee JH, Lee CK, Suh CH, Kang HS, Hong CP, Choi SN. Serum concentrations of per- and poly-fluoroalkyl substances and factors associated with exposure in the general adult population in South Korea. *Int J Hyg Environ Health* 2017;220:1046-54. DOI PubMed
103. Lindh CH, Rylander L, Toft G, et al. Blood serum concentrations of perfluorinated compounds in men from Greenlandic Inuit and European populations. *Chemosphere* 2012;88:1269-75. DOI PubMed
104. Liu Y, Su J, van Dam RM, et al. Dietary predictors and plasma concentrations of perfluorinated alkyl acids in a Singapore population. *Chemosphere* 2017;171:617-24. DOI PubMed
105. Yamaguchi M, Arisawa K, Uemura H, et al. Consumption of seafood, serum liver enzymes, and blood levels of PFOS and PFOA in the Japanese population. *J Occup Health* 2013;55:184-94. DOI PubMed
106. Beesoon S, Genuis SJ, Benskin JP, Martin JW. Exceptionally high serum concentrations of perfluorohexanesulfonate in a Canadian family are linked to home carpet treatment applications. *Environ Sci Technol* 2012;46:12960-7. DOI PubMed
107. Martin JW, Asher BJ, Beesoon S, Benskin JP, Ross MS. PFOS or PreFOS? Are perfluorooctane sulfonate precursors (PreFOS) important determinants of human and environmental perfluorooctane sulfonate (PFOS) exposure? *J Environ Monit* 2010;12:1979-2004. DOI PubMed
108. Strynar MJ, Lindstrom AB. Perfluorinated compounds in house dust from Ohio and North Carolina, USA. *Environ Sci Technol* 2008;42:3751-6. DOI PubMed
109. Jian JM, Chen D, Han FJ, et al. A short review on human exposure to and tissue distribution of per- and polyfluoroalkyl substances (PFASs). *Sci Total Environ* 2018;636:1058-69. DOI PubMed
110. Cuffney M, Wilkie A, Kotlarz N, et al. Factors associated with per-and polyfluoroalkyl substances (PFAS) serum concentrations in residents of Wilmington, North Carolina: the genx exposure study. *SSRN* 2023. DOI
111. Colles A, Bruckers L, Den Hond E, et al. Perfluorinated substances in the Flemish population (Belgium): levels and determinants of variability in exposure. *Chemosphere* 2020;242:125250. DOI PubMed
112. Raudhatunnisa T, Wilantika N. Performance comparison of hot-deck imputation, K-nearest neighbor imputation, and predictive mean matching in missing value handling, case study: March 2019 SUSENAS Kor dataset. *icdsos* 2022;2021:753-70. DOI
113. Olsen GW, Burris JM, Ehresman DJ, et al. Half-life of serum elimination of perfluorooctanesulfonate,perfluorohexanesulfonate, and perfluorooctanoate in retired fluorochemical production workers. *Environ Health Perspect* 2007;115:1298-305. DOI PubMed PMC