Methods

Indicator

B7. Perfluorochemicals in women ages 16 to 49 years: Median concentrations in blood serum, 1999-March 2020.

Summary

Since the 1970s, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Health and Nutrition Examination Surveys (NHANES), a series of U.S. national surveys of the health and nutrition status of the non-institutionalized civilian population. The National Center for Environmental Health at CDC measures environmental chemicals in blood and urine samples collected from NHANES participants.[i] This indicator uses blood serum perfluorochemical (PFC) concentrations of four PFCs: perfluorooctane sulfonic acid (PFOS), perfluorooctanoic acid (PFOA), perfluorohexane sulfonic acid (PFHxS), and perfluorononanoic acid (PFNA). The NHANES biennial survey cycles (1999-2000, 2003-2020) included blood serum PFC data for children and adults ages 12 years and over.[ii] For 2013-March 2020, the linear and branched isomers of PFOA and PFOS were separately measured and were summed for these analyses.

Indicator B7 is the median blood serum concentration for each of these PFCs for women ages 16 to 49 years, stratified by survey period. For 1999-2016, survey periods are two years based on the duration of each NHANES survey cycle. The 2017-2020 survey cycles include data up to March 2020 and are grouped as a single survey period with a duration of 3.2 years[iii]. The median is the estimated concentration such that 50 percent of all non-institutionalized civilian women ages 16 to 49 years have a PFC concentration below this level; the population distribution was adjusted by age-specific birth rates to estimate the median prenatal exposure to PFCs.

Supplementary Tables: Table B7a presents the 95th percentile concentrations of each of these PFCs for women ages 16 to 49 years, stratified by survey period. The 95th percentile is the estimated concentration such that 95 percent of all non-institutionalized civilian women ages 16 to 49 years have a PFC concentration below this level. Table B7b presents the median concentrations of each of these PFCs for women ages 16 to 49 years for 2015-March 2020, stratified by race/ethnicity and family income. Table B7c presents the 95th percentile concentrations of each of these PFCs for women ages 16 to 49 years for 2015-March 2020, stratified by race/ethnicity and family income. The survey data were weighted to account for over-sampling, non-response, and non-coverage.

[i] Centers for Disease Control and Prevention. 2009. Fourth National Report on Human Exposure to Environmental Chemicals. Atlanta, GA. Available at: www.cdc.gov/exposurereport.
[ii] Blood serum data from NHANES 2001-2002 are not included in Indicator B7 because the data were pooled and thus not comparable to data from other years.
[iii] The NHANES 2019-2020 survey was prematurely suspended in March 2020 and did not produce nationally representative sample data. The NHANES 2017-March 2020 data set is analyzed herein.

Data Summary

Indicator B7. Perfluorochemicals in women ages 16 to 49 years: Median concentrations in blood serum, 1999-March 2020.

Data		Blood Serum PFC for four PFCs.								
PFOS	Years	1999-2000	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014†	2015-2016†	2017-March 2020†
	Limits of Detection (ng/mL)*	0.2	0.4	0.2	0.2	0.2	0.2	0.1	0.1	0.1
	Number of Values	444	577	684	556	653	542	612	585	851
	Number of Non-missing Values**	444 (100%)	504 (87%)	626 (92%)	495 (89%)	610 (93%)	490 (90%)	569 (93%)	534 (91%)	768 (90%)
	Number of Missing Values**	0 (0%)	73 (13%)	58 (8%)	61 (11%)	43 (7%)	52 (10%)	43 (7%)	51 (9%)	83 (10%)
	Percentage Below Limit of Detection***	0	0	0	0	0	0	n-PFOS: 1 Sm-PFOS: 2	n-PFOS: 2 Sm-PFOS: 3	n-PFOS: 0 Sm-PFOS: 1
PFOA	Years	1999-2000	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014†	2015-2016†	2017-March 2020†
	Limits of Detection (ng/mL)*	0.2	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Number of Values	444	577	684	556	653	542	612	585	851
	Number of Non-missing Values**	444 (100%)	504 (87%)	626 (92%)	495 (89%)	610 (93%)	490 (90%)	569 (93%)	534 (91%)	768 (90%)
	Number of Missing Values**	0 (0%)	73 (13%)	58 (8%)	61 (11%)	43 (7%)	52 (10%)	43 (7%)	51 (9%)	83 (10%)
	Percentage Below Limit of Detection***	0	1	0	0	0	0	n-PFOA: 1 Sb-PFOA: 88	n-PFOA: 1 Sb-PFOA: 100	n-PFOA: 0 Sb-PFOA: 97

PFHxS	Years	1999-2000	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014	2015-2016	2017-March 2020
	Limits of Detection (ng/mL)*	0.1	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Number of Values	444	577	684	556	653	542	612	585	851
	Number of Non-missing Values**	444 (100%)	504 (87%)	626 (92%)	495 (89%)	610 (93%)	490 (90%)	571 (93%)	534 (91%)	768 (90%)
	Number of Missing Values**	0 (0%)	73 (13%)	58 (8%)	61 (11%)	43 (7%)	52 (10%)	41 (7%)	51 (9%)	83 (10%)
	Percentage Below Limit of Detection***	0	4	7	1	0	1	1	2	0
PFNA	Years	1999-2000	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014	2015-2016	2017-March 2020
	Limits of Detection (ng/mL)*	0.2	0.1	0.1	0.082	0.082	0.08	0.1	0.1	0.1
	Number of Values	444	577	684	556	653	542	612	585	851
	Number of Non-missing Values**	444 (100%)	504 (87%)	626 (92%)	495 (89%)	610 (93%)	490 (90%)	571 (93%)	534 (91%)	768 (90%)
	Number of Missing Values**	0 (0%)	73 (13%)	58 (8%)	61 (11%)	43 (7%)	52 (10%)	41 (7%)	51 (9%)	83 (10%)
	Percentage Below Limit of Detection***	7	2	1	0	1	0	1	3	1

* The Limit of Detection (LOD) is defined as the level at which the measurement has a 95% probability of being greater than zero.

**Non-missing values include those below the analytical LOD, which are reported as LOD/√2. Missing values are the number of sampled women ages 16 to 49 years in the Mobile Examination Center (MEC) sub-sample that have no value reported for the particular variable used in calculating the indicator.

***This percentage is survey-weighted using the NHANES MEC survey weights for the given period and is weighted by age-specific birth rates.

† For 2013-March 2020, the linear and branched isomers of PFOA were measured separately (linear = n-PFOA, branched = Sb-PFOA) and the linear and branched isomers of PFOS were measured separately (linear = n-PFOS, branched = Sm-PFOS). All four isomers had the same limit of detection and the same numbers of missing and non-missing measurements.

National Health and Nutrition Examination Surveys (NHANES)

Since the 1970s, the National Center for Health Statistics, a division of the Centers for Disease Control and Prevention, has conducted the National Health and Nutrition Examination Surveys (NHANES), a series of U.S. national surveys of the health and nutrition status of the non-institutionalized civilian population. The National Center for Environmental Health at CDC measures environmental chemicals in blood and urine samples collected from NHANES participants.

This indicator uses blood serum concentrations of four PFCs from NHANES 1999-2000, 2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, and 2017-March 2020 in women ages 16 to 49. The NHANES data were obtained from the NHANES website: https://wwwn.cdc.gov/nchs/nhanes/default.aspx. Following the CDC recommended approach, values below the analytical limit of detection (LOD) were replaced by LOD/√2.[iv]

This analysis uses the four PFCs and four isomers listed in the following table.

PFC Abbreviation	Full name	SAS name (1999-2000)	SAS name for non-detect comment code (1999-2000)*	SAS name (2003-March 2020)	SAS name for non-detect comment code (2003-March 2020)*
PFOS**	Perfluorooctane sulfonic acid	SPFOS	SPFOSLC	LBXPFOS	LBDPFOSL
PFOA**	Perfluorooctanoic acid	SPFOA	SPFOALC	LBXPFOA	LBDPFOAL
PFHxS	Perfluorohexane sulfonic acid	SPFHS	SPFHSLC	LBXPFHS	LBSPFHSL
PFNA	Perfluorononanoic acid	SPFNA	SPFNALC	LBXPFNA	LBDPFNAL
n-PFOS**	Linear perfluorooctane sulfonate			SSNPFOS, LBXNFOS	SDNPFOSL, LBDNFOSL
Sm-PFOS**	Monomethyl branched isomers of PFOS			SSMPFOS, LBXMFOS	SDMPFOSL, LBDMFOSL
n-PFOA**	Linear perfluorooctanoate			SSNPFOA, LBXNFOA	SDNPFOAL, LBDNFOAL
Sb-PFOA**	Branched isomers of perfluorooctanoate			SSBPFOA, LBXBFOA	SDBPFOAL, LBDBFOAL

*The non-detect comment code equals 1 if the measurement is below the analytical limit of detection and equals 0 if the measurement is at or above the analytical limit of detection.
**PFOS and PFOA were measured in 1999-2012. The isomers n-PFOS, Sm-PFOS, n-PFOA, and Sb-PFOA were measured in 2013-March 2020.

For 2013-March 2020, PFOS is calculated by summing the measurements of the isomers n-PFOS and Sm-PFOS, and PFOA is calculated by summing the measurements of the isomers n-PFOA and Sb-PFOA. For PFOS, each sample measurement either had both of the isomers or neither isomer. For PFOA, each sample measurement either had both of the isomers or neither isomer. The total PFOS is treated as being below the analytical limit of detection if one or both of the constituent isomers is below the analytical limit of detection. The total PFOA is treated as being

[iv] See Hornung RW, Reed LD. 1990. Estimation of average concentration in the presence of nondetectable values. *Applied Occupational and Environmental Hygiene* 5:46–51.

below the analytical limit of detection if one or both of the constituent isomers is below the analytical limit of detection. For calculating the sums, any constituent isomer that is below its analytical limit of detection (LOD) is replaced by LOD/√2 before adding the concentrations.

The NHANES use a complex multi-stage, stratified, clustered sampling design. Certain demographic groups were deliberately over-sampled, including Mexican-Americans, Blacks, and, from 2007 onwards, All Hispanics, then, from 2011 onwards, Asians, to increase the reliability and precision of estimates of health status indicators for these population subgroups. The publicly released data includes survey weights to adjust for the over-sampling, non-response, and non-coverage. The statistical analyses used the applicable MEC survey weights (WTMEC2YR for1999-2000, WTSA2YR for 2003-2006 and 2011-2012, WTSC2YR for 2007-2010, WTSB2YR for 2013-2016, and WTSBAPRP for 2017-March 2020) to re-adjust the blood serum PFC data to represent the national population.

Age-Specific Birth Rates

In addition to the NHANES MEC survey weights, the data for women ages 16 to 49 were also weighted by the birth rate for women of the given age and race/ethnicity to estimate prenatal exposures. Thus, the overall weight in each two-year period is the product of the NHANES MEC survey weight and the total number of births in the survey period for the given age and race/ethnicity, divided by twice the corresponding population of women at the midpoint of the survey period:[v]

Adjusted Survey Weight =
MEC survey weight × U.S. Births (NHANES cycle, age, race/ethnicity) /
{Number of years in NHANES cycle × U.S. Women (NHANES cycle midpoint, age, race/ethnicity)}.

All NHANES cycle data are based on a 2 year survey period with a midpoint of 1 year, except for the 2017-March 2020 pre-pandemic cycle. NHANES 2017-March 2020 covers a 3.2 year survey period with a midpoint of 1.6 years.[vi]

Race/Ethnicity and Family Income

For Tables B7b and B7c the percentiles were calculated for demographic strata defined by the combination of race/ethnicity and family income.

The family income was characterized based on the INDFMPIR variable, which is the ratio of the family income to the poverty level. The National Center for Health Statistics used the U.S. Census Bureau Current Population Survey definition of a "family" as "a group of two people or more (one of whom is the householder) related by birth, marriage, or adoption and residing together" to group household members into family units, and the corresponding family income

[v] Axelrad, D.A., Cohen, J. 2011. Calculating summary statistics for population chemical biomonitoring in women of child-bearing age with adjustment for age-specific natality. *Environmental Research* 111 (1) 149-155.
[vi] Akinbami L.J. et al. 2022. National Health and Nutrition Examination Survey, 2017-March 2020 prepandemic file: Sample design, estimation, and analytic guidelines. *National Center for Health Statistics. Vital Health Stat* 2(190).

for the respondent was obtained during the interview. The U.S. Census Bureau defines annual poverty level money thresholds varying by family size and composition. The poverty income ratio (PIR) is the family income divided by the poverty level for that family. Family income was stratified into the following groups:

- Below Poverty Level: PIR < 1
- Above Poverty Level: PIR ≥ 1
- Unknown Income: PIR is missing

For the 5.2-year period 2015-March 2020, the weighted percentage of women ages 16 to 49 years with unknown income was 9%.

Race/ethnicity was characterized using the RIDRETH1 variable. The possible values of this variable are:

- 1. Mexican American
- 2. Other Hispanic
- 3. Non-Hispanic White
- 4. Non-Hispanic Black
- 5. Other Race – Including Multi-racial
- "." Missing

Category 5 includes: all non-Hispanic single race responses other than White or Black; and multi-racial responses.

For this indicator, the RIDRETH1 categories 2, 5, and missing were combined into a single "All Other Races/Ethnicities" category. This produced the following categories:

- White non-Hispanic: RIDRETH1 = 3
- Black non-Hispanic: RIDRETH1 = 4
- Mexican-American: RIDRETH1 = 1
- All Other Races/Ethnicities: RIDRETH1 = 2 or 5 or missing

The "All Other Races/Ethnicities" category includes multiracial persons and individuals whose racial or ethnic identity is not White non-Hispanic, Black non-Hispanic, or Mexican-American. Except for non-Mexican-American Hispanics in 2007-March 2020 and Asian non-Hispanics in 2011-March 2020, persons of "All Other Races/Ethnicities" are selected into the survey with a probability that is very much lower than White non-Hispanic, Black non-Hispanic and Mexican-American individuals, and as a group they are not representative of all other race and ethnicities in the United States.

Calculation of Indicator

Indicator B7 is the median for blood serum PFC in women of ages 16 to 49 years, stratified by survey period. The median is the estimated concentration such that 50 percent of all non-institutionalized civilian women ages 16 to 49 years have blood serum PFC concentrations below

this level. To adjust the NHANES data to represent prenatal exposures, the data for each woman surveyed was multiplied by the estimated number of births per woman of the given age and race/ethnicity. Table B7a presents the 95th percentile for blood serum PFC in women of ages 16 to 49 years, stratified by survey period. The 95th percentile is the estimated concentration such that 95 percent of all non-institutionalized civilian women ages 16 to 49 years have blood serum PFC concentrations below this level. Table B7b presents the median for blood serum PFC in women of ages 16 to 49 years, stratified both by race/ethnicity and family income. Table B7c presents the 95th percentile for blood serum PFC in women of ages 16 to 49 years, stratified both by race/ethnicity and family income.

To simply demonstrate the calculations, we will use the NHANES 2007-2008 blood serum PFOS values for women ages 16 to 49 years of all race/ethnicities and all incomes as an example. We have rounded all the numbers to make the calculations easier:

We begin with all the non-missing NHANES 2007-2008 blood serum PFOS values for women ages 16 to 49 years. Assume for the sake of simplicity that valid data on blood serum PFOS were available for every sampled woman. Each sampled woman has an associated annual survey weight, WTSCYR, that estimates the annual number of U.S. women represented by that sampled woman. Each sampled woman also has an associated birth rate giving the numbers of annual births per woman of the given age, race, and ethnicity. The product of the annual survey weight and the birth rate estimates the annual number of U.S. births represented by that sampled woman, which we will refer to as the adjusted survey weight. For example, the lowest blood serum PFOS measurement for a woman between 16 and 49 years of age is 0.4 ng/mL with an annual survey weight of 230,000, a birth rate of 0.004, and thus an adjusted survey weight of 1,000, and so represents 1,000 births. The total of the adjusted survey weights for the sampled women equals 4 million, the total number of annual U.S. births to women ages 16 to 49 years. The second-lowest measurement is 0.5 ng/mL with an adjusted survey weight of 100, and so represents another 100 U.S. births. The highest measurement was 93 ng/mL, with an adjusted survey weight of 20, and so represents another 20 U.S. births.

To calculate the median, we can use the adjusted survey weights to expand the data to the entire U.S. population of births to women ages 16 to 49. We have 1,000 values of 0.4 ng/mL from the lowest measurement, 100 values of 0.5 ng/mL from the second lowest measurement, and so on, up to 20 values of 93 ng/mL from the highest measurement. Arranging these 4 million values in increasing order, the 2 millionth value is 8.7 ng/mL. Since half of the values are below 8.8 and half of the values are above 8.7, the median equals 8.7 ng/mL. To calculate the 95th percentile, as in Table B7a, note that 95 percent of 4 million equals 3.8 million. The 3.8 millionth value is 22.8 ng/mL. Since 95 percent of the values are below 22.8, the 95th percentile equals 22.8 ng/mL

The calculations need to take into account that blood serum PFOS measurements were not available for every respondent, and to use exact rather than rounded numbers. There were blood serum PFOS measurements for only 495 of the 556 sampled women ages 16 to 49 years. The adjusted survey weights for all 556 sampled women add up to 4.1 million, the U.S. population of births to women ages 16 to 49. The adjusted survey weights for the 495 sampled women with blood serum PFOS data add up to 3.8 million. Thus, the available data represent 3.8 million values and so represent only 92% of the U.S. population of births. The median and 95th percentiles are given by the 1.90 millionth (50% of 3.8 million) and 3.61 millionth (95% of 3.8

million) U.S. birth's value. These calculations assume that the sampled women with valid blood serum PFOS data are representative of women giving birth without valid blood serum PFOS data. The calculations also assume that the sampled women are representative of women that actually gave birth in 2007-2008, since NHANES information on pregnancy and births was not incorporated into the analysis.

Equations

These percentile calculations can also be given as the following mathematical equations, which are based on the default percentile calculation formulas from Statistical Analysis System (SAS) software. Exclude all missing blood serum PFOS values. Suppose there are n women of ages 16 to 49 years with valid blood serum PFOS values. Arrange the blood serum PFC concentrations in increasing order (including tied values) so that the lowest concentration is x(1) with an adjusted survey weight of w(1), the second lowest concentration is x(2) with an adjusted survey weight of w(2), …, and the highest concentration is x(n) with an adjusted survey weight of w(n).

1. Sum all the adjusted survey weights to get the total weight W:

$$W = \Sigma[1 \le i \le n]\ w(i)$$

2. Find the largest number i so that the total of the weights for the i lowest values is less than or equal to W/2.

$$\Sigma[j \le i]\ w(j) \le W/2 < \Sigma[j \le i + 1]\ w(j)$$

3. Calculate the median using the results of the second step. We either have

$$\Sigma[j \le i]\ w(j) = W/2 < \Sigma[j \le i + 1]\ w(j)$$

or

$$\Sigma[j \le i]\ w(j) < W/2 < \Sigma[j \le i + 1]\ w(j)$$

In the first case we define the median as the average of the i'th and i + 1'th values:

$$\text{Median} = [x(i) + x(i + 1)]/2 \text{ if } \Sigma[j \le i]\ w(j) = W/2$$

In the second case we define the median as the i + 1'th value:

$$\text{Median} = x(i + 1) \text{ if } \Sigma[j \le i]\ w(j) < W/2$$

(The estimated median does not depend upon how the tied values of x(j) are ordered).

A similar calculation applies to the 95th percentile. The first step to calculate the sum of the weights, W, is the same. In the second step, find the largest number i so that the total of the weights for the i lowest values is less than or equal to 0.95W.

$$\Sigma[j \leq i]\ w(j) \leq 0.95W < \Sigma[j \leq i + 1]\ w(j)$$

In the third step we calculate the 95th percentile using the results of the second step. We either have

$$\Sigma[j \leq i]\ w(j) = 0.95W < \Sigma[j \leq i + 1]\ w(j)$$

or

$$\Sigma[j \leq i]\ w(j) < 0.95W < \Sigma[j \leq i + 1]\ w(j)$$

In the first case we define the 95th percentile as the average of the i'th and i + 1'th values:

$$95^{th}\ \text{Percentile} = [x(i) + x(i + 1)]/2 \text{ if } \Sigma[j \leq i]\ w(j) = 0.95W$$

In the second case we define the 95th percentile as the i + 1'th value:

$$95^{th}\ \text{Percentile} = x(i + 1) \text{ if } \Sigma[j \leq i]\ w(j) < 0.95W$$

Relative Standard Error

The uncertainties of the median and 95th percentile values were calculated using a revised version of the CDC method given in CDC 2005,[vii] Appendix C, and the SAS® program provided by CDC. The method uses the Clopper-Pearson binomial confidence intervals adapted for complex surveys by Korn and Graubard (see Korn and Graubard, 1999,[viii] p. 65). The following text is a revised version of the Appendix C. For the birth rate adjusted calculations for women ages 16 to 49, the sample weight is adjusted by multiplying by the age-specific birth rate.

Step 1: Use SAS® Proc Univariate to obtain a point estimate P_{SAS} of the percentile value. Use the Weight option to assign the exact correct sample weight for each chemical result.

Step 2: Use SUDAAN® Proc Descript with Taylor Linearization DESIGN = WR (i.e., sampling with replacement) and the proper sampling weight to estimate the proportion (p) of subjects with results less than and not equal to the percentile estimate P_{SAS} obtained in Step 1 and to obtain the standard error (se_p) associated with this proportion estimate. Compute the degrees-of-freedom adjusted effective sample size

$$n_{df} = (t_{num}/t_{denom})^2\ p(1 - p)/(se_p^2)$$

where t_{num} and t_{denom} are 0.975 critical values of the Student's t distribution with degrees of freedom equal to the sample size minus 1 and the number of PSUs minus the number of strata, respectively. Note: the degrees of freedom for t_{denom} can vary with the demographic sub-group of interest.

Step 3: After obtaining an estimate of p (i.e., the proportion obtained in Step 2), compute the Clopper-Pearson 95% confidence interval ($P_L(x,n_{df})$, $P_U(x,n_{df})$) as follows:

$$P_L(x,n_{df}) = v_1F_{v1,v2}\ (0.025)/(v_2 + v_1F_{v1,v2}(0.025))$$
$$P_U(x,n_{df}) = v_3F_{v3,v4}\ (0.975)/(v_4 + v_3F_{v3,v4}(0.975))$$

[vii] CDC Third National Report on Human Exposure to Environmental Chemicals. 2005
[viii] Korn E. L., Graubard B. I. 1999. *Analysis of Health Surveys.* Wiley.

where x is equal to p times n_{df}, $v_1 = 2x$, $v_2 = 2(n_{df} - x + 1)$, $v_3 = 2(x + 1)$, $v_4 = 2(n_{df} - x)$, and $F_{d1,d2}(\beta)$ is the β quantile of an F distribution with d1 and d2 degrees of freedom. (Note: If n_{df} is greater than the actual sample size or if p is equal to zero, then the actual sample size should be used.) This step will produce a lower and an upper limit for the estimated proportion obtained in Step 2.

Step 4: Use SAS Proc Univariate (again using the Weight option to assign weights) to determine the chemical percentile values P_{CDC}, L_{CDC} and U_{CDC} that correspond to the proportion p obtained in Step 2 and its lower and upper limits obtained in Step 3. Do not round the values of p and the lower and upper limits. For example, if p = 0.4832, then P_{CDC} is the 48.32'th percentile value of the chemical. The alternative percentile estimates P_{CDC} and P_{SAS} are not necessarily equal.

Step 5: Use the confidence interval from Step 4 to estimate the standard error of the estimated percentile P_{CDC}:

$$\text{Standard Error } (P_{CDC}) = (U_{CDC} - L_{CDC}) / (2t_{denom})$$

Step 6: Use the estimated percentile P_{CDC} and the standard error from Step 4 to estimate the relative standard error of the estimated percentile P_{CDC}:

$$\text{Relative Standard Error (\%)} = [\text{Standard Error } (P_{CDC}) / P_{CDC}] \times 100\ \%$$

The tabulated estimated percentile is the value of P_{SAS} given in Step 1. The relative standard error is given in Step 6, using P_{CDC} and its standard error.

The relative standard error depends upon the survey design. For this purpose, the public release version of NHANES includes the variables SDMVSTRA and SDMVPSU, which are the Masked Variance Unit pseudo-stratum and pseudo-primary sampling unit (pseudo-PSU). For approximate variance estimation, the survey design can be approximated as being a stratified random sample with replacement of the pseudo-PSUs from each pseudo-stratum; the true stratum and PSU variables are not provided in the public release version to protect confidentiality. If the relative standard error is too high, then the estimated percentile will not be accurately estimated. Furthermore, if the degrees of freedom (from Step 2) is too low, then the relative standard error will be less accurately estimated and thus may be underestimated. For these reasons, percentiles with high relative standard errors or with low degrees of freedom are unstable or unreliable.

Percentiles with a relative standard error less than 30% and with 12 or more degrees of freedom were treated as being reliable and were tabulated. Percentiles with a relative standard error that is 30% or greater but less than 40% and with 12 or more degrees of freedom were treated as being unstable; these values were tabulated but were flagged to be interpreted with caution. Percentiles with a relative standard error less than 40% and with between 7 and 11 degrees of freedom were also treated as being unstable; these values were tabulated but were flagged to be interpreted with caution. Percentiles with a relative standard error that is 40% or greater, or without an estimated relative standard error, or with 6 or less degrees of freedom, were treated as being unreliable; these values were not tabulated and were flagged as having a large uncertainty.

Questions and Comments

Questions regarding these methods, and suggestions to improve the description of the methods, are welcome. Please use the "Contact Us" link at the bottom of any page in the America's Children and the Environment website.

Statistical Comparisons

Statistical analyses of the percentiles were used to determine whether the differences between percentiles for different demographic groups were statistically significant. For these analyses, the percentiles and their standard errors were calculated for each combination of age group, income group (below poverty, at or above poverty, unknown income), and race/ethnicity group using the method described in the "Relative Standard Error" section. In the notation of that section, the percentile and standard error are the values of P_{CDC} and Standard Error (P_{CDC}), respectively. These calculated standard errors account for the survey weighting and design and, for women, for the age-specific birth rate.

Using a weighted linear regression model, the percentile was assumed to be the sum of explanatory terms for age, income and/or race/ethnicity and a random error term; the error terms were assumed to be approximately independent and normally distributed with a mean of zero and a variance equal to the square of the standard error. In this model, the weight is the inverse of the variance, so that percentiles with larger standard errors are given less of a statistical weight in the fitted regression model. Using this model, the difference in the value of a percentile between different demographic groups is statistically significant if the difference between the corresponding sums of explanatory terms is statistically significantly different from zero. A p-value at or below 0.05 implies that the difference is statistically significant at the 5% significance level. No adjustment is made for multiple comparisons.

For each type of comparison, we present unadjusted and adjusted analyses. The unadjusted analyses directly compare a percentile between different demographic groups. The adjusted analyses add other demographic explanatory variables to the statistical model and use the statistical model to account for the possible confounding effects of these other demographic variables. For example, the unadjusted race/ethnicity comparisons use and compare the percentiles between different race/ethnicity pairs. The adjusted race/ethnicity comparisons use the percentiles for each age/ income/race/ethnicity combination. The adjusted analyses add age, and income terms to the statistical model and compare the percentiles between different race/ethnicity pairs after accounting for the effects of the other demographic variables. For example, if White non-Hispanics tend to have higher family incomes than Black non-Hispanics, and if the blood perfluorochemical level strongly depends on family income only, then the unadjusted differences between these two race/ethnicity groups would be significant but the adjusted difference (taking into account income) would not be significant.

Comparisons between pairs of race/ethnicity groups and between income groups are shown in Tables 1 and 2, respectively, for women ages 16 to 49 years. In Table 1, for the unadjusted "All incomes" comparisons, the only explanatory variables are terms for each race/ethnicity group. For these unadjusted comparisons, the statistical tests compare the percentiles for each pair of race/ethnicity groups. For the adjusted "All incomes (adjusted for age, income)" comparisons, the explanatory variables are terms for each race/ethnicity group, together with terms for each age and income group. For these adjusted comparisons, the statistical test compares the pair of race/ethnicity groups after accounting for any differences in the age and income distributions between the race/ethnicity groups.

In Table 1, for the unadjusted "Below Poverty Level" and "At or Above Poverty Level" comparisons, the only explanatory variables are terms for each of the twelve race/ethnicity/income combinations (combinations of four race/ethnicity groups and three income groups). For example, in row 1, the p-value for "Below Poverty Level" compares White non-Hispanics below the poverty level with Black non-Hispanics below the poverty level. The same set of explanatory variables are used in Table 2 for the unadjusted comparisons between one race/ethnicity group below the poverty level and the same race/ethnicity group at or above the poverty level. The corresponding adjusted analyses include extra explanatory variables for age, so that race/ethnicity/income groups are compared after accounting for any differences due to age. Although these comparisons only involve the two income groups with known incomes, these statistical models were fitted to all three income groups (including those with unknown income) to make a more general, better fitting model; this approach has no impact on the unadjusted p-values but has a small impact on the adjusted p-values. Also in Table 2, the unadjusted p-value for the population "All" compares the percentiles for women ages 16 to 49 years below poverty level with those at or above poverty level, using the explanatory variables for the two income groups (below poverty, at or above poverty), excluding those with unknown income. The adjusted p-value includes adjustment terms for age and race/ethnicity in the model.

Additional comparisons are shown in Table 3 for women ages 16 to 49 years. Comparisons are shown for differences between those below poverty and those at or above poverty, and for changes over time (trends). The Against = "income" unadjusted p-value compares the percentiles for those below poverty level with those at or above poverty level, using the explanatory variables for the two income groups (below poverty, at or above poverty). The adjusted p-value includes adjustment terms for age and race/ethnicity in the model. The Against = "year" p-value examines whether the linear trend in the percentiles is statistically significant (using the percentiles for each NHANES period regressed against the midpoint of that period); the adjusted model for trend adjusts for demographic changes in the populations from year to year by including terms for age, income, and race/ethnicity.

For women, the age groups used were 16-19, 20-24, 25-29, 30-39, and 40-49.

For more details on these statistical analyses, see the memorandum by Cohen (2010).[ix]

[ix] Cohen, J. 2010. *Selected statistical methods for testing for trends and comparing years or demographic groups in ACE NHIS and NHANES indicators.* Memorandum submitted to Dan Axelrad, EPA, 21 March, 2010.

Table 1. Statistical significance tests comparing the percentiles of PFCs in women ages 16 to 49 years, between pairs of race/ethnicity groups, for 2015-March 2020.

				P-VALUES					
Variable	Percentile	First race/ethnicity group	Second race/ethnicity group*	All incomes	All incomes (adjusted for age, income)	Below Poverty Level	Below Poverty Level (adjusted for age)	At or Above Poverty Level	At or Above Poverty Level (adjusted for age)
PFOS	50	White non-Hispanic	Black non-Hispanic	0.023	< 0.001	0.076	< 0.001	0.071	0.126
PFOS	50	White non-Hispanic	Mexican-American	< 0.001	< 0.001	0.007	< 0.001	< 0.001	< 0.001
PFOS	50	White non-Hispanic	Other	0.032	0.117	0.171	0.273	0.200	0.045
PFOS	50	Black non-Hispanic	Mexican-American	0.014	0.174	0.265	0.134	0.006	0.045
PFOS	50	Black non-Hispanic	Other	1.000	0.018	0.519	0.147	0.718	0.843
PFOS	50	Mexican-American	Other	0.032	< 0.001	0.044	0.526	0.007	< 0.001
PFOA	50	White non-Hispanic	Black non-Hispanic	< 0.001	0.011	0.395	0.026	0.006	0.098
PFOA	50	White non-Hispanic	Mexican-American	< 0.001	0.105	0.347	0.006	0.001	< 0.001
PFOA	50	White non-Hispanic	Other	0.525	0.919	0.896	0.778	0.571	0.634
PFOA	50	Black non-Hispanic	Mexican-American	1.000	0.497	1.000	< 0.001	0.291	0.104
PFOA	50	Black non-Hispanic	Other	0.017	0.123	0.610	0.753	0.195	0.088
PFOA	50	Mexican-American	Other	0.017	0.319	0.589	0.221	0.079	0.002
PFHxS	50	White non-Hispanic	Black non-Hispanic	0.010	< 0.001	0.120	< 0.001	0.003	0.002
PFHxS	50	White non-Hispanic	Mexican-American	< 0.001	< 0.001	0.387	< 0.001	< 0.001	< 0.001
PFHxS	50	White non-Hispanic	Other	0.198	< 0.001	0.756	< 0.001	0.010	0.016
PFHxS	50	Black non-Hispanic	Mexican-American	0.004	0.418	0.169	0.071	1.000	0.941
PFHxS	50	Black non-Hispanic	Other	0.004	0.031	< 0.001	0.152	0.187	0.312
PFHxS	50	Mexican-American	Other	< 0.001	0.178	0.168	0.998	0.003	0.166
PFNA	50	White non-Hispanic	Black non-Hispanic	< 0.001	< 0.001	1.000	< 0.001	0.144	0.074
PFNA	50	White non-Hispanic	Mexican-American	0.004	< 0.001	0.257	0.014	0.018	< 0.001
PFNA	50	White non-Hispanic	Other	0.004	< 0.001	0.002	0.002	0.068	0.465
PFNA	50	Black non-Hispanic	Mexican-American	0.004	0.034	0.207	0.412	0.233	< 0.001
PFNA	50	Black non-Hispanic	Other	0.004	< 0.001	< 0.001	< 0.001	1.000	0.290
PFNA	50	Mexican-American	Other	1.000	< 0.001	0.259	< 0.001	0.171	< 0.001
PFOS	95	White non-Hispanic	Black non-Hispanic	0.952	0.164	0.855	0.006	0.957	< 0.001
PFOS	95	White non-Hispanic	Mexican-American	0.199	< 0.001	< 0.001	< 0.001	0.262	< 0.001
PFOS	95	White non-Hispanic	Other	0.794	< 0.001	0.002	0.081	0.704	0.790
PFOS	95	Black non-Hispanic	Mexican-American	0.206	< 0.001	0.018	< 0.001	0.279	0.030
PFOS	95	Black non-Hispanic	Other	0.763	< 0.001	0.003	0.024	0.762	0.001

				P-VALUES					
Variable	**Percentile**	**First race/ethnicity group**	**Second race/ethnicity group***	**All incomes**	**All incomes (adjusted for age, income)**	**Below Poverty Level**	**Below Poverty Level (adjusted for age)**	**At or Above Poverty Level**	**At or Above Poverty Level (adjusted for age)**
PFOS	95	Mexican-American	Other	0.230	< 0.001	< 0.001	0.993	0.327	< 0.001
PFOA	95	White non-Hispanic	Black non-Hispanic	0.054	< 0.001	0.633	< 0.001	0.018	< 0.001
PFOA	95	White non-Hispanic	Mexican-American	< 0.001	< 0.001	0.712	< 0.001	0.003	< 0.001
PFOA	95	White non-Hispanic	Other	0.250	< 0.001	0.283	0.720	0.071	0.098
PFOA	95	Black non-Hispanic	Mexican-American	0.141	< 0.001	0.826	0.836	0.622	0.141
PFOA	95	Black non-Hispanic	Other	0.029	< 0.001	0.702	0.248	0.005	< 0.001
PFOA	95	Mexican-American	Other	0.004	< 0.001	0.468	0.234	0.003	< 0.001
PFHxS	95	White non-Hispanic	Black non-Hispanic	0.207	< 0.001	0.117	< 0.001	0.165	0.402
PFHxS	95	White non-Hispanic	Mexican-American	0.185	< 0.001	0.427	0.650	0.209	0.002
PFHxS	95	White non-Hispanic	Other	0.137	< 0.001	0.280	< 0.001	0.078	< 0.001
PFHxS	95	Black non-Hispanic	Mexican-American	0.792	0.660	0.892	< 0.001	0.860	0.001
PFHxS	95	Black non-Hispanic	Other	0.362	0.028	0.948	0.002	0.352	< 0.001
PFHxS	95	Mexican-American	Other	0.584	0.011	0.940	< 0.001	0.347	< 0.001
PFNA	95	White non-Hispanic	Black non-Hispanic	0.745	< 0.001	0.220	< 0.001	0.359	0.028
PFNA	95	White non-Hispanic	Mexican-American	1.000	< 0.001	0.007	< 0.001	0.662	< 0.001
PFNA	95	White non-Hispanic	Other	0.150	< 0.001	0.003	0.396	0.149	0.714
PFNA	95	Black non-Hispanic	Mexican-American	0.792	< 0.001	0.886	< 0.001	1.000	< 0.001
PFNA	95	Black non-Hispanic	Other	0.745	< 0.001	0.543	0.047	0.006	0.006
PFNA	95	Mexican-American	Other	0.444	< 0.001	0.279	< 0.001	0.189	< 0.001

* "Other" represents the "All Other Races/Ethnicities" category, which includes all other races and ethnicities not specified, together with those individuals who report more than one race.

Table 2. Statistical significance tests comparing the percentiles of PFCs in women ages 16 to 49 years, between those below poverty level and those at or above poverty level, for 2015-March 2020.

			P-Values for difference between income levels	
Variable	**Percentile**	**Population***	**Unadjusted**	**Adjusted (for age)****
PFOS	50	All	0.188	< 0.001
PFOS	50	White non-Hispanic	1.000	0.181
PFOS	50	Black non-Hispanic	0.293	0.447
PFOS	50	Mexican-American	1.000	< 0.001
PFOS	50	Other	0.508	0.496
PFOA	50	All	< 0.001	0.524
PFOA	50	White non-Hispanic	0.108	0.005
PFOA	50	Black non-Hispanic	0.235	0.035
PFOA	50	Mexican-American	0.559	< 0.001
PFOA	50	Other	0.395	0.114
PFHxS	50	All	1.000	0.002
PFHxS	50	White non-Hispanic	0.392	0.425
PFHxS	50	Black non-Hispanic	0.184	0.009
PFHxS	50	Mexican-American	1.000	0.053
PFHxS	50	Other	1.000	0.017
PFNA	50	All	0.004	0.356
PFNA	50	White non-Hispanic	0.003	0.795
PFNA	50	Black non-Hispanic	0.060	0.029
PFNA	50	Mexican-American	0.329	0.053
PFNA	50	Other	0.055	0.001
PFOS	95	All	0.501	< 0.001
PFOS	95	White non-Hispanic	0.690	0.208
PFOS	95	Black non-Hispanic	0.686	< 0.001
PFOS	95	Mexican-American	0.809	< 0.001
PFOS	95	Other	0.002	0.069
PFOA	95	All	< 0.001	< 0.001
PFOA	95	White non-Hispanic	< 0.001	< 0.001
PFOA	95	Black non-Hispanic	0.815	0.028
PFOA	95	Mexican-American	0.697	0.065
PFOA	95	Other	0.014	0.006
PFHxS	95	All	0.232	< 0.001
PFHxS	95	White non-Hispanic	0.957	0.036
PFHxS	95	Black non-Hispanic	0.045	< 0.001
PFHxS	95	Mexican-American	0.753	< 0.001
PFHxS	95	Other	0.795	0.243
PFNA	95	All	0.595	< 0.001
PFNA	95	White non-Hispanic	0.001	< 0.001
PFNA	95	Black non-Hispanic	0.442	0.047
PFNA	95	Mexican-American	0.220	< 0.001
PFNA	95	Other	0.185	0.009

* "Other" represents the "All Other Races/Ethnicities" category, which includes all other races and ethnicities not specified, together with those individuals who report more than one race.
** Comparison for "All" is adjusted for age and race/ethnicity; comparisons for race/ethnicity categories are adjusted for age.

Table 3. Other statistical significance tests comparing the percentiles of PFCs in women ages 16 to 49 years, for 2015-March 2020 (trends for 1999-March 2020).

					P-VALUES	
Variable	**Percentile**	**From**	**To**	**Against**	**Unadjusted**	**Adjusted***
PFOS	50	2015	March 2020	income	0.188	< 0.001
PFOS	50	1999	March 2020	year	< 0.001	< 0.001
PFOA	50	2015	March 2020	income	< 0.001	0.524
PFOA	50	1999	March 2020	year	< 0.001	< 0.001
PFHxS	50	2015	March 2020	income	1.000	0.002
PFHxS	50	1999	March 2020	year	< 0.001	< 0.001
PFNA	50	2015	March 2020	income	0.004	0.356
PFNA	50	1999	March 2020	year	< 0.001	< 0.001
PFOS	95	2015	March 2020	income	0.501	< 0.001
PFOS	95	1999	March 2020	year	< 0.001	< 0.001
PFOA	95	2015	March 2020	income	< 0.001	< 0.001
PFOA	95	1999	March 2020	year	< 0.001	< 0.001
PFHxS	95	2015	March 2020	income	0.232	< 0.001
PFHxS	95	1999	March 2020	year	< 0.001	< 0.001
PFNA	95	2015	March 2020	income	0.595	< 0.001
PFNA	95	1999	March 2020	year	< 0.001	< 0.001

*For Against = "income," the comparison is between those below the poverty level and those at or above the poverty level, and the p-values are adjusted for age and race/ethnicity.
For Against = "year" the comparison is the trend over different years, and the p-values are adjusted for age, race/ethnicity, and income.

Data Files

The following files are needed to calculate this indicator. The files together with the survey documentation and SAS programs for reading in the data are available at the NHANES website: http://www.cdc.gov/nchs/nhanes.htm.

- NHANES 1999-2000: Demographic file demo.xpt. Surplus Specimen Laboratory Component: Perflouroalkyl Chemicals (Surplus Sera) Laboratory file sspfc_a.xpt. The demographic file demo.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA), pseudo-PSU (SDMVPSU), and two-year Mobile Examination Center (MEC) weight (WTMEC2YR). The Perfluoroalkyl Chemicals laboratory file sspfc_a.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (SPFOS, SPFOA, SPFHS, SPFNA), and the PFC non-detect comment codes (SPFOSLC, SPFOALC, SPFHSLC, SPFNALC). The two files are merged using the common variable SEQN.

- NHANES 2003-2004: Demographic file demo_c.xpt. Polyfluoroalkyl Chemicals Laboratory file l24pfc_c.xpt. The demographic file demo_c.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum

(SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Polyfluoroalkyl Chemicals laboratory file l24pfc_c.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (LBXPFOS, LBXPFOA, LBXPFHS, LBXPFNA), the PFC non-detect comment codes (LBDPFOSL, LBDPFOAL, LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample A (WTSA2YR). The two files are merged using the common variable SEQN.

- NHANES 2005-2006: Demographic file demo_d.xpt. Polyfluoroalkyl Chemicals Laboratory file pfc_d.xpt. The demographic file demo_d.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Polyfluoroalkyl Chemicals laboratory file pfc_d.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (LBXPFOS, LBXPFOA, LBXPFHS, LBXPFNA), the PFC non-detect comment codes (LBDPFOSL, LBDPFOAL, LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample A (WTSA2YR). The two files are merged using the common variable SEQN.

- NHANES 2007-2008: Demographic file demo_e.xpt. Polyfluoroalkyl Chemicals Laboratory file pfc_e.xpt. The demographic file demo_e.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Polyfluoroalkyl Chemicals laboratory file pfc_e.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (LBXPFOS, LBXPFOA, LBXPFHS, LBXPFNA), the PFC non-detect comment codes (LBDPFOSL, LBDPFOAL, LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample C (WTSC2YR). The two files are merged using the common variable SEQN.

- NHANES 2009-2010: Demographic file demo_f.xpt. Polyfluoroalkyl Chemicals Laboratory file pfc_f.xpt. The demographic file demo_f.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Polyfluoroalkyl Chemicals laboratory file pfc_f.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (LBXPFOS, LBXPFOA, LBXPFHS, LBXPFNA), the PFC non-detect comment codes (LBDPFOSL, LBDPFOAL, LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample C (WTSC2YR). The two files are merged using the common variable SEQN.

- NHANES 2011-2012: Demographic file demo_g.xpt. Polyfluoroalkyl Chemicals Laboratory file pfc_g.xpt. The demographic file demo_g.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Polyfluoroalkyl Chemicals laboratory file pfc_g.xpt contains SEQN, the PFCs PFOS, PFOA, PFHxS and PFNA (LBXPFOS, LBXPFOA, LBXPFHS, LBXPFNA), the PFC non-detect comment codes

(LBDPFOSL, LBDPFOAL, LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample A (WTSCAYR). The two files are merged using the common variable SEQN.

- NHANES 2013-2014: Demographic file demo_h.xpt. Perfluoroalkyl and Polyfluoroalkyl Substances Laboratory file pfas_h.xpt. Perfluoroalkyl and Polyfluoroalkyl Substances – Linear and Branched PFOS and PFOA Isomers (Surplus) Laboratory file sspfas_h.xpt The demographic file demo_h.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Perfluoroalkyl and Polyfluoroalkyl Substances laboratory file pfas_h.xpt contains SEQN, PFCs PFHxS and PFNA (LBXPFHS, LBXPFNA), the PFC non-detect comment codes (LBDPFHSL, LBDPFNAL), and the two-year MEC weight of sub-sample B (WTSCBYR). The Perfluoroalkyl and Polyfluoroalkyl Substances – Linear and Branched PFOS and PFOA Isomers (Surplus) laboratory file sspfas_h.xpt contains SEQN, linear perfluorooctane sulfonate (n-PFOS, SAS name SSNPFOS), monomethyl branched isomers of PFOS (Sm-PFOS, SAS name SSMPFOS), linear perfluorooctanoate (n-PFOA, SAS name SSNPFOA), branched isomers of perfluorooctanoate (Sb-PFOA, SAS name SSBPFOA), the corresponding PFC non-detect comment codes (SDNPFOSL, SDMPFOSL, SDNPFOAL, SDBPFOAL), and the two-year MEC weight of sub-sample B (WTSCBYR). The three files are merged using the common variable SEQN.

- NHANES 2015-2016: Demographic file demo_i.xpt. Perfluoroalkyl and Polyfluoroalkyl Laboratory file pfas_i.xpt. The demographic file demo_i.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Perfluoroalkyl and Polyfluoroalkyl laboratory file pfas_i.xpt contains SEQN, PFCs perflourohexane sulfonic acid (PFHxS, SAS name LBXPFHS), perflourononanoic acid (PFNA, SAS name LBXPFNA), linear perfluorooctane sulfonate (n-PFOS, SAS name LBXNFOS), monomethyl branched isomers of PFOS (Sm-PFOS, SAS name LBXMFOS), linear perfluorooctanoate (n-PFOA, SAS name LBXNFOA), branched isomers of perfluorooctanoate (Sb-PFOA, SAS name LBXBFOA), the corresponding PFC non-detect comment codes (LBDPFHSL, LBDPFNAL, LBDNFOSL, LBDMFOSL, LBDNFOAL, LBDBFOAL), and the two-year MEC weight of sub-sample B (WTSCBYR). The two files are merged using the common variable SEQN.

- NHANES 2017-March 2020: Demographic file P_DEMO.xpt. Perfluoroalkyl and Polyfluoroalkyl Laboratory fileP_PFAS.xpt. The demographic file P_DEMO.xpt is a SAS transport file that contains the subject identifier (SEQN), age (RIDAGEYR), sex (RIAGENDR), race/ethnicity (RIDRETH1), poverty income ratio (INDFMPIR), pseudo-stratum (SDMVSTRA) and the pseudo-PSU (SDMVPSU). The Perfluoroalkyl and Polyfluoroalkyl laboratory file P_PFAS.xpt contains SEQN, PFCs perflourohexane sulfonic acid (PFHxS, SAS name LBXPFHS), perflourononanoic acid (PFNA, SAS name LBXPFNA), linear perfluorooctane sulfonate (n-PFOS, SAS name LBXNFOS), monomethyl branched isomers of PFOS (Sm-PFOS, SAS name LBXMFOS), linear perfluorooctanoate (n-PFOA, SAS name LBXNFOA), branched isomers of

perfluorooctanoate (Sb-PFOA, SAS name LBXBFOA), the corresponding PFC non-detect comment codes (LBDPFHSL, LBDPFNAL, LBDNFOSL, LBDMFOSL, LBDNFOAL, LBDBFOAL), and the 3.2 year MEC weight of sub-sample BA (WTSBAPRP). The two files are merged using the common variable SEQN.