# Methods

## Indicator

B12. Perchlorate in women ages 16 to 49 years: Median and 95th percentile concentrations in urine, 2001-March 2020

B13. Perchlorate in children ages 6 to 17 years: Median and 95th percentile concentrations in urine, 2001-March 2020 (NEW, previously Table B13c)[i]

## 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 noninstitutionalized civilian population. The National Center for Environmental Health at CDC measures environmental chemicals in blood and urine samples collected from NHANES participants.[ii] These indicators use urine measurements of perchlorate in women ages 16 to 49 years and children ages 6 to 17 years. The NHANES 2001-2002, 2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, 2013-2014, 2015-2016, and 2017-March 2020 surveys included urine perchlorate data for children and adults ages 6 years and over.

Indicator B12 is the median and 95th percentile concentrations of perchlorate in the population of non-institutionalized civilian women ages 16 to 49 years for 2001-March 2020. Indicator B13 (NEW, previously Table B13c) is the median and 95th percentile concentrations of perchlorate in the population of children ages 6 to 17 years for 2001-March 2020.[i] The median is the estimated concentration such that 50% of the population have a perchlorate concentration below this level; for indicator B12, the population distribution was adjusted by age-specific birth rates to reflect exposures to women who are pregnant or may become pregnant. The 95th percentile is the estimated concentration such that 95% of the population have a perchlorate concentration below this level. Table B12a presents the median concentration of perchlorate for women ages 16 to 49 years for 2015-March 2020, stratified both by race/ethnicity and family income. Table B12b presents the 95th percentile concentration of perchlorate for women ages 16 to 49 years for 2015-March 2020, stratified both by race/ethnicity and family income. Table B13a presents the median concentration of perchlorate for children ages 6 to 17 years in 2015-March 2020, stratified both by race/ethnicity and family income. Table B13b presents the 95th percentile concentration of perchlorate for children ages 6 to 17 years in 2015-March 2020, stratified both by race/ethnicity and family income. Table B13c presents the median and 95th percentile concentrations of perchlorate for children ages 6 to 17 in 2015-March 2020, stratified by age. The survey data were weighted to account for over-sampling, non-response, and non-coverage.

[i] Beginning October 2025, perchlorate in children ages 6 to 17 years is reported as a unique indicator, Indicator B13. Perchlorate in children ages 6 to 17 years was previously reported as Supplementary Table B13c.
[ii] Centers for Disease Control and Prevention. 2009. Fourth National Report on Human Exposure to Environmental Chemicals. Atlanta, GA. Available at: www.cdc.gov/exposurereport.

**Data Summary**

| Indicator | Indicator B12. Perchlorate in women ages 16 to 49 years: Median and 95th percentile concentrations in urine, 2001-March 2020. | | | |
|---|---|---|---|---|
| Time Period | 2001-March 2020 | | | |
| Data | Urine perchlorate in women ages 16 to 49 years | | | |
| Years | 2001-2002 | 2003-2004 | 2005-2006 | 2007-2008 |
| Limits of Detection (µg/L)* | 0.05 | 0.05 | 0.05 | 0.05 |
| Number of values | 700 | 623 | 2085 | 1749 |
| Number of Non-missing Values** | 657 (94%) | 616 (99%) | 1921 (92%) | 1608 (92%) |
| Number of Missing Values** | 43 (6%) | 7 (1%) | 164 (8%) | 141 (8%) |
| Percentage Below Limit of Detection*** | 0 | 0 | 0 | 0 |
| Years | 2009-2010 | 2011-2012 | 2013-2014 | 2015-2016 |
| Limits of Detection (µg/L)* | 0.05 | 0.05 | 0.05 | 0.05 |
| Number of values | 686 | 542 | 632 | 587 |
| Number of Non-missing Values** | 662 (97%) | 528 (97%) | 610 (97%) | 563 (96%) |
| Number of Missing Values** | 24 (3%) | 14 (3%) | 22 (3%) | 24 (3%) |
| Percentage Below Limit of Detection*** | 0 | 0 | 0 | |
| Years | 2017-March 2020 | | | |
| Limits of Detection (µg/L)* | 0.05 | | | |
| Number of values | 844 | | | |
| Number of Non-missing Values** | 821 (97%) | | | |
| Number of Missing Values** | 23 (3%) | | | |
| Percentage Below Limit of Detection*** | 0 | | | |

* 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.

| Indicator | Indicator B13. Perchlorate in children ages 6 to 17 years: Median and 95th percentile concentrations in urine, 2001-March 2020 | | | |
|---|---|---|---|---|
| Time Period | 2001-March 2020 | | | |
| Data | Urine perchlorate in children ages 6 to 17 years. | | | |
| Years | 2001-2002 | 2003-2004 | 2005-2006 | 2007-2008 |
| Limits of Detection (µg/L)* | 0.05 | 0.05 | 0.05 | 0.05 |
| Number of values | 1058 | 859 | 2849 | 2207 |
| Number of Non-missing Values** | 1021 (97%) | 858 (100%) | 2626 (92%) | 2012 (91%) |
| Number of Missing Values** | 37 (3%) | 1 (0%) | 223 (8%) | 195 (9%) |
| Percentage Below Limit of Detection*** | 0 | 0 | 0 | 0 |
| Years | 2009-2010 | 2011-2012 | 2013-2014 | 2015-2016 |
| Limits of Detection (µg/L)* | 0.05 | 0.05 | 0.05 | 0.05 |
| Number of values | 769 | 706 | 775 | 750 |
| Number of Non-missing Values** | 734 (95%) | 683 (97%) | 728 (94%) | 692 (92%) |
| Number of Missing Values** | 35 (5%) | 23 (3%) | 47 (6%) | 58 (8%) |
| Percentage Below Limit of Detection*** | | | | 0 |
| Years | 2017-March 2020 | | | |
| Limits of Detection (µg/L)* | 0.05 | | | |
| Number of values | 1073 | | | |
| Number of Non-missing Values** | 1022 (95%) | | | |
| Number of Missing Values** | 51 (5%) | | | |
| Percentage Below Limit of Detection*** | 0 | | | |

* 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 children ages 6 to 17 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.

**Overview of 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/nhanes_questionnaires.htm.

- NHANES 2001-2002: Demographic file demo_b.xpt. Nitrate, thiocyanate, perchlorate (Surplus Urine) laboratory file ssno3p_b.xpt. The demographic file demo_b.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 Nitrate, thiocyanate, perchlorate (Surplus Urine) laboratory file ssno3p_b.xpt contains SEQN, urine perchlorate (SSXUP8), and the two-year Mobile Examination Center (MEC) sub-sample weight (WTUIO2YR). The two files are merged using the common variable SEQN.

- NHANES 2003-2004: Demographic file demo_c.xpt. Urinary Perchlorate laboratory file l04per_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 Urinary Perchlorate laboratory file l04per_c.xpt contains SEQN, urine perchlorate (URXUP8), and the two-year MEC sub-sample C weight (WTSC2YR). The two files are merged using the common variable SEQN.

- NHANES 2005-2006: Demographic file demo_d.xpt. Urinary Nitrate, Urinary Perchlorate, Urinary Thiocyanate laboratory file pernt_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), pseudo-PSU (SDMVPSU), and the two-year MEC weight (WTMEC2YR). The Urinary Nitrate, Urinary Perchlorate, Urinary Thiocyanate laboratory file pernt_d.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2007-2008: Demographic file demo_e.xpt. Urinary Nitrate, Urinary Perchlorate, Urinary Thiocyanate laboratory file pernt_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) , pseudo-PSU (SDMVPSU), and the two year MEC weight (WTMEC2YR). The Urinary Nitrate, Urinary Perchlorate, Urinary Thiocyanate laboratory file pernt_e.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2009-2010: Demographic file demo_f.xpt. Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_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), pseudo-PSU (SDMVPSU), and the two-year MEC sub-sample weight (WTSA2YR). Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_f.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment

code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2011-2012: Demographic file demo_f.xpt. Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_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), pseudo-PSU (SDMVPSU), and the two-year MEC subsample weight (WTSA2YR). Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_g.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2013-2014: Demographic file demo_h.xpt. Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_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), pseudo-PSU (SDMVPSU), and the two-year MEC sub-sample weight (WTSA2YR). Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_h.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2015-2016: Demographic file demo_i.xpt. Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_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), pseudo-PSU (SDMVPSU), and the two-year MEC sub-sample weight (WTSA2YR). Perchlorate, Nitrate & Thiocyanate – Urine laboratory file pernt_i.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

- NHANES 2017-March 2020: Demographic file P_DEMO.xpt. Perchlorate, Nitrate & Thiocyanate – Urine laboratory file P_PERNT.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), pseudo-PSU (SDMVPSU), and the 3.2-year MEC sub-sample weight (WTSAPRP). Perchlorate, Nitrate & Thiocyanate – Urine laboratory file P_PERNT.xpt contains SEQN, urine perchlorate (URXUP8), and the perchlorate non-detect comment code (URDUP8LC). The two files are merged using the common variable SEQN.

**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 noninstitutionalized 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 urine perchlorate measurements from NHANES 2001-2002, 2003-2004, 2005-2006, 2007-2008. 2009-2010, 2011-2012, 2013-2014, 2015-2016, and 2017-March 2020 in women ages 16 to 49 and children ages 6 to 17. 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.[iii]

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, 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 sample or sub-sample survey weights (WTUIO2YR for 2001-2002, WTSC2YR for 2003-2004, WTMEC2YR for 2005-2006 and 2007-2008, WTSA2YR for 2009-2010, 2011-2012, 2013-2014, and 2015-2016, WTSAPRP for 2017-March 2020) to re-adjust the urine perchlorate data to represent the national population.

**Age-Specific Birth Rates**

In addition to the NHANES MEC survey weights, the data for women of child-bearing age (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 survey period is the product of the NHANES survey weight and the total number of births in the two calendar years for the given age and race/ethnicity, divided by twice the corresponding population of women at the midpoint of the survey period:[iv]

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.[v]

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

**Race/Ethnicity and Family Income**

For these indicators, 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 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 years period 2015-March 2020, the weighted percentage of women ages 16 to 49 years with unknown income was 12% and the weighted percentage of children ages 6 to 17 years with unknown income was 8%.

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 2007-2014 and Asian non-Hispanics in 2015-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 B12 is the median and 95$^{th}$ percentile for urine perchlorate in the population of noninstitutionalized civilian women of ages 16 to 49 years, stratified by NHANES survey cycle. Indicator B13 is the median and 95$^{th}$ percentile for urine perchlorate in the population of children of ages 6 to 17 years, stratified by NHANES survey cycle. The median for each population is the estimated concentration such that 50% of the population of interest have urine perchlorate concentrations below this level. The 95$^{th}$ percentile for each population is the estimated concentration such that 95% of the population have urine perchlorate 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 B12a presents the median for urine perchlorate in women of ages 16 to 49 years in 2015-March 2020, stratified by race/ethnicity and family income. Table B12b presents the 95$^{th}$ percentile for urine perchlorate in women of ages 16 to 49 years in 2015-March 2020, stratified by race/ethnicity and family income. Table B13a presents the median for urine perchlorate in children of ages 6 to 17 years in 2015-March 2020, stratified by race/ethnicity and family income. The birth rate adjustment was not applied to children ages 6 to 17. Table B13b presents the 95$^{th}$ percentile for urine perchlorate in children of ages 6 to 17 years in 2015-March 2020, stratified by race/ethnicity and family income. Table B13c presents the median and 95$^{th}$ percentile concentration of perchlorate for children ages 6 to 17 in 2015-March 2020, stratified by age.

To simply demonstrate the calculations, we will use the NHANES 2007-2008 urine perchlorate 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 urine perchlorate values for women ages 16 to 49 years. Assume for the sake of simplicity that valid perchlorate data were available for every sampled woman. Each sampled woman has an associated annual survey weight that estimates the annual number of U.S. women represented by that sampled woman. For 2007-2008, the associated annual survey weight for each woman is defined as WTMEC2YR. 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 urine perchlorate measurement for a woman between 16 and 49 years of age is 0.1 μg/L with an annual survey weight of 25,000, a birth rate of 0.008, and thus an adjusted survey weight of 200, and so represents 200 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 also 0.1 µg/L with an adjusted survey weight of 8, and so represents another 8 U.S. births. The highest measurement is 220 µg/L with an adjusted survey weight of 30, and so represents another 30 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 200 values of 0.1 µg/L from the lowest measurement, 8 values of 0.1 µg/L from the second lowest measurement, and so on, up to 30 values of 220 µg/L from the highest measurement. Arranging these 4 million values in increasing order, the 2 millionth value is 3.4 µg/L. Since half of the values are below 3.4 and half of the values are above 3.4, the median equals 3.4 µg/L. To calculate the 95$^{th}$ percentile, note that 95% of 4 million equals 3.8 million. The 3.8 millionth value is 16.5 µg/L. Since 95% of the values are below 16.5, the 95$^{th}$ percentile equals 16.5 µg/L.

In reality, the calculations need to take into account that urine perchlorate measurements were not available for every respondent, and to use exact rather than rounded numbers. There were urine perchlorate measurements for only 1608 of the 1749 sampled women ages 16 to 49 years. The adjusted survey weights for all 1749 sampled women add up to 4.2 million, the U.S. population of births to women ages 16 to 49. The adjusted survey weights for the 1608 sampled women with urine perchlorate data add up to 3.9 million. Thus the available data represent 3.9 million values and so represent only 94% of the U.S. population of births. The median and 95$^{th}$ percentiles are given by the 1.95 millionth (50% of 3.9 million) and 3.7 millionth (95% of 3.9 million) U.S. birth's value. These calculations assume that the sampled women with valid urine perchlorate data are representative of women giving birth without valid urine perchlorate 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 urine perchlorate values. Suppose there are n women of ages 16 to 49 years with valid urine perchlorate values. Arrange the urine perchlorate 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:

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

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.

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

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

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

or

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

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

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

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

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

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

A similar calculation applies to the 95$^{th}$ 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 95$^{th}$ 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 95$^{th}$ 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 95$^{th}$ 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,[vi] 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,[vii] 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:

$$P_L(x,n_{df}) = v_1 F_{v1,v2}(0.025)/(v_2 + v_1 F_{v1,v2}(0.025))$$
$$P_U(x,n_{df}) = v_3 F_{v3,v4}(0.975)/(v_4 + v_3 F_{v3,v4}(0.975))$$

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 $P_{CDC}$:

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

Relative Standard Error (%) = [Standard Error ($P_{CDC}$) / $P_{CDC}$] × 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 across time (trend) or for different demographic groups were statistically significant. For these analyses, the percentiles and their standard errors were calculated for each combination of age group, sex (in the cases of children), 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, sex, 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 percentiles across time or 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/sex/income/race/ethnicity combination. The adjusted analyses add age, sex, 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 urine perchlorate 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 of percentiles across time (trend) are shown in Table 1 for women ages 16 to 49 years and in Table 4 for children ages 6 to 17 years. 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 accounts for demographic changes in the populations from year to year by including terms for age, sex (for children), income, and race/ethnicity. Table 4 includes additional comparisons between children's age groups; the adjusted model accounts for sex, race/ethnicity, and income.

Comparisons between pairs of race/ethnicity groups are shown in Table 2 for women ages 16 to 49 years and in Table 5 for children ages 6 to 17 years. Comparisons between income groups are shown in Table 3 for women ages 16 to 49 years and in Table 6 for children ages 6 to 17 years. In Tables 2 and 5, 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, sex, income)" comparisons, the explanatory variables are terms for each race/ethnicity group together with terms for each age, sex (for children), 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, sex (for children) and income distributions between the race/ethnicity groups. The adjustment for sex is applicable only for children, and thus appears only in Tables 4, 5 and 6.

In Tables 2 and 5, 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 Tables 3 and 6 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 and sex (for children), so that race/ethnicity/income groups are compared after accounting for any differences due to age or sex. 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 p-values from unadjusted models but has a small impact on the p-values from adjusted models. Also in Tables 3 and 6, the p-value for the population "All" compares the percentiles for women ages 16 to 49 years or children ages 6 to 17 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 model includes terms for age, sex (for children), and race/ethnicity in the model.

For women, the age groups used were 16-19, 20-24, 25-29, 30-39, and 40-49. For children, the age groups used were 6-10, 11-15, and 16-17. For more details on these statistical analyses, see the memorandum by Cohen (2010).[viii]

Table 7 shows the results of post-hoc analyses comparing percentiles between children ages 6-17 and women ages 16-49. Two-sided Welch's t-test was performed using $P_{CDC}$, standard errors, and effective sample sizes for the two age groups, derived using the approach described in the "Relative Standard Error" section. Welch's t-test assumes groups are independent; overlapping groups can impact estimates of variability and should be considered when interpreting results. In this case, the groups overlap for ages 16-17. Table B13c presents the median and 95$^{th}$ percentile concentrations for the 16-17 age group.

[viii] 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 perchlorate in women ages 16 to 49 years across time for 2001-March 2020.

| | | | | | P-VALUES | |
|---|---|---|---|---|---|---|
| **Variable** | **Percentile** | **From** | **To** | **Against** | **Unadjusted** | **Adjusted*** |
| Perchlorate | 50 | 2001 | March 2020 | year | < 0.001 | < 0.001 |
| Perchlorate | 95 | 2001 | March 2020 | year | < 0.001 | < 0.001 |

* For Against = "year", the model was adjusted for age, race/ethnicity, and income. The comparison is the trend over different years.

Table 2. Statistical significance tests comparing the percentiles of perchlorate 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)**** |
| Perchlorate | 50 | White non-Hispanic | Black non-Hispanic | 1.000 | 0.079 | 0.989 | 0.962 | 0.451 | 0.074 |
| Perchlorate | 50 | White non-Hispanic | Mexican-American | 0.856 | < 0.001 | 0.850 | 0.001 | 0.788 | 0.530 |
| Perchlorate | 50 | White non-Hispanic | Other | 0.760 | 0.422 | 0.844 | 0.014 | 0.865 | 0.797 |
| Perchlorate | 50 | Black non-Hispanic | Mexican-American | 0.844 | 0.049 | 0.844 | 0.011 | 0.814 | 0.021 |
| Perchlorate | 50 | Black non-Hispanic | Other | 0.724 | 0.397 | 0.836 | 0.028 | 0.524 | 0.020 |
| Perchlorate | 50 | Mexican-American | Other | 0.960 | 0.008 | 0.991 | 0.894 | 0.879 | 0.640 |
| Perchlorate | 95 | White non-Hispanic | Black non-Hispanic | 0.171 | 0.043 | 0.870 | 0.171 | 0.412 | 0.331 |
| Perchlorate | 95 | White non-Hispanic | Mexican-American | 0.147 | < 0.001 | 0.996 | < 0.001 | 0.513 | 0.070 |
| Perchlorate | 95 | White non-Hispanic | Other | 0.468 | 0.001 | 0.431 | < 0.001 | 0.713 | 0.967 |
| Perchlorate | 95 | Black non-Hispanic | Mexican-American | 0.696 | < 0.001 | 0.877 | 0.062 | 0.959 | < 0.001 |
| Perchlorate | 95 | Black non-Hispanic | Other | 0.664 | < 0.001 | 0.666 | < 0.001 | 0.537 | 0.259 |
| Perchlorate | 95 | Mexican-American | Other | 0.478 | < 0.001 | 0.602 | 0.028 | 0.657 | 0.015 |

* "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.
** P-values are based on models that were adjusted for demographic variables.

Table 3. Statistical significance tests comparing the percentiles of perchlorate 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**** |
| Perchlorate | 50 | All | 0.710 | 0.196 |
| Perchlorate | 50 | White non-Hispanic | 0.860 | 0.803 |
| Perchlorate | 50 | Black non-Hispanic | 0.527 | 0.062 |
| Perchlorate | 50 | Mexican-American | 0.930 | 0.004 |
| Perchlorate | 50 | Other | 0.987 | 0.016 |
| Perchlorate | 95 | All | 0.571 | 0.041 |
| Perchlorate | 95 | White non-Hispanic | 0.905 | 0.394 |
| Perchlorate | 95 | Black non-Hispanic | 0.824 | 0.002 |
| Perchlorate | 95 | Mexican-American | 0.507 | 0.530 |
| Perchlorate | 95 | Other | 0.146 | 0.002 |

* "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.
** For Population="All", the model is adjusted for age and race/ethnicity; for comparisons within race/ethnicity categories the model is adjusted for age.

Table 4. Statistical significance tests comparing the percentiles of perchlorate in children 6 to 17 years across time for 2015-March 2020 and between age groups for 2015-March 2020 .

| | | | | | P-VALUES | |
|---|---|---|---|---|---|---|
| **Variable** | **Percentile** | **From** | **To** | **Against*** | **Unadjusted** | **Adjusted**** |
| Perchlorate | 50 | 2015 | March 2020 | age | < 0.001 | < 0.001 |
| Perchlorate | 50 | 2001 | March 2020 | year | < 0.001 | < 0.001 |
| Perchlorate | 95 | 2015 | March 2020 | age | 0.019 | < 0.001 |
| Perchlorate | 95 | 2001 | March 2020 | year | < 0.001 | < 0.001 |

* For Against = "year", the comparison is the trend over different years. For Against = "age," the comparison is between the age groups 6-10, 11-15, and 16-17 years.
** P-values are based on models adjusted for sex, race/ethnicity, and income. For Against="year" the model is also adjusted for age group.

Table 5. Statistical significance tests comparing the percentiles of perchlorate in children ages 6 to 17 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)** |
| Perchlorate | 50 | White non-Hispanic | Black non-Hispanic | 0.369 | 0.232 | 0.988 | 0.547 | 0.389 | 0.544 |
| Perchlorate | 50 | White non-Hispanic | Mexican-American | 0.700 | 0.833 | 0.863 | 0.398 | 0.932 | 0.151 |
| Perchlorate | 50 | White non-Hispanic | Other | 0.265 | 0.219 | 0.623 | 0.368 | 0.972 | 0.507 |
| Perchlorate | 50 | Black non-Hispanic | Mexican-American | 0.713 | 0.272 | 0.842 | 0.634 | 0.528 | 0.031 |
| Perchlorate | 50 | Black non-Hispanic | Other | 0.890 | 0.001 | 0.463 | < 0.001 | 0.396 | 0.879 |
| Perchlorate | 50 | Mexican-American | Other | 0.617 | 0.066 | 0.786 | 0.002 | 0.954 | 0.054 |
| Perchlorate | 95 | White non-Hispanic | Black non-Hispanic | 0.888 | < 0.001 | 0.708 | < 0.001 | 0.663 | 0.370 |
| Perchlorate | 95 | White non-Hispanic | Mexican-American | 0.056 | 0.778 | 0.753 | 0.006 | 0.102 | 0.001 |
| Perchlorate | 95 | White non-Hispanic | Other | 0.656 | < 0.001 | 0.900 | 0.008 | 0.562 | 0.824 |
| Perchlorate | 95 | Black non-Hispanic | Mexican-American | 0.039 | < 0.001 | 1.000 | 0.192 | 0.060 | < 0.001 |
| Perchlorate | 95 | Black non-Hispanic | Other | 0.549 | < 0.001 | 0.536 | < 0.001 | 0.356 | 0.141 |
| Perchlorate | 95 | Mexican-American | Other | 0.139 | < 0.001 | 0.627 | < 0.001 | 0.260 | < 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.
** P-values are based on models that were adjusted for demographic variables.

Table 6. Statistical significance tests comparing the percentiles of perchlorate in children ages 6 to 17 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** |
| Perchlorate | 50 | All | 0.230 | 0.101 |
| Perchlorate | 50 | White non-Hispanic | 0.589 | 0.518 |
| Perchlorate | 50 | Black non-Hispanic | 0.811 | 0.471 |
| Perchlorate | 50 | Mexican-American | 0.489 | 0.103 |
| Perchlorate | 50 | Other | 0.053 | 0.001 |
| Perchlorate | 95 | All | 0.518 | 0.243 |
| Perchlorate | 95 | White non-Hispanic | 0.704 | 0.007 |
| Perchlorate | 95 | Black non-Hispanic | 0.735 | 0.002 |
| Perchlorate | 95 | Mexican-American | 0.257 | < 0.001 |
| Perchlorate | 95 | Other | 0.304 | < 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.
** For Population="All" the model is adjusted for age, sex, and race/ethnicity; for comparisons within race/ethnicity categories the model is adjusted for age and sex.

Table 7. Statistical significance tests comparing the percentiles of perchlorate between children ages 6 to 17 years and women ages 16 to 49 years, for 2015-March 2020.

| | | | | | Comparison of Age Groups* | | | |
|---|---|---|---|---|---|---|---|---|
| **Age Group** | **Percentile** | **$P_{CDC}$**** | **Standard Error($P_{CDC}$)**** | **Effective Sample Size**** | **Mean Difference (95% CI)** | **T Statistic** | **Degrees of Freedom** | **P-Values** |
| Children, 3-17 | 50 | 2.9 | 0.114 | 1714 | | | | |
| Women, 16-49 | 50 | 2.18 | 0.156 | 1384 | 0.72 (0.342, 1.098) | 3.7309 | 2644 | <0.001 |
| Children, 3-17 | 95 | 10.9 | 0.512 | 1714 | | | | |
| Women, 16-49 | 95 | 9.0 | 1.024 | 1384 | 1.9 (-0.346, 4.146 | 1.569 | 2057.14 | 0.097 |

* Statistical tests assume independence for dependent age groups.
** Statistical tests were performed using chemical percentile estimates ($P_{CDC}$), standard error of $P_{CDC}$, and degrees-of-freedom adjusted effective sample sizes from the revised CDC method for relative standard errors.