[2. Methods](#_yqwbau70ycz1)

[2.1 Study Population](#_wuz13j21hufv)

[2.2 Data](#_fgg5qnimlk7)

[2.3 Exposure assessment via urinary concentrations](#_wuz13j21hufv)

[2.4 Thyroid function measures](#_wuz13j21hufv)

[2.5 Covariate Selection](#_u44cfiyhcagu)

[2.6 Statistical Analysis](#_lnb4azi0mg25)

# 2. Methods

## 2.1 Study Population

Participant data in this study come from the National Health and Nutrition Examination Survey (NHANES). NHANES is a nationally-representative cross-sectional survey of the United States civilian population gathered yearly using multistage area probability sampling (Johnson et al., 2014). Sample weights were used in this study to account for oversampling and establish comparability with estimated population distributions. NHANES data are gathered annually but are released in two-year cycles; this study analyzed the datasets corresponding to two year pairs: 2013-2014 and 2015-2016.

For our analyses, we included a population restricted to participants above the age of 20 who had no missing values for any study variables of interest.

## 2.2 Data

*2.2.2 LOD*

Upon preliminary exploratory analysis of the data trends and missingness, imputation of values under the limit of detection (LOD) were carried out. Reported detection limits in 100 μL of urine were as follows: BPA, 0.2 ng/mL; BPF, 0.2 ng/mL; BPS, 0.1 ng/mL. observations that fell below this value were calculated as the limit of detection divided by the square root of 2 (Hornung and Reed 1990).All biomarker measurements in urine were adjusted for individual specific creatinine clearance.

*2.2.3 Transformation of exposure variables for specific analyses (logistic, nonlinear, quantile)*

To fit quantile regression models, we log transformed BPA, BPF, and BPS exposure measurements to account for normality assumption. For principal component analysis (PCA), we centered and scaled each urinary measurement of environmental chemical exposure.

*2.2.4 NHANES Survey Weights*

To account for the NHANES complex survey design, potential oversampling, survey non-response, and post-stratification, all analyses were conducted with survey weight-adjusted data.

## 2.3 Exposure assessment via urinary concentrations

Data on urinary concentrations of chemicals and metabolites were gathered from a one-third sample of the total NHANES study population. Included in this study are Bisphenol A (BPA), Bisphenol S (BPS, 4,4′-sulfonyldiphenol), and Bisphenol F (BPF, 4,4′-dihydroxydiphenyl methane); methyl, ethyl, propyl, and butyl parabens; 2,4- and 2,5-dichlorophenol; triclocarban; and triclosan.

Urinary concentrations were gathered using a method based on previous research (Zhou et al, 2014). This method entailed “on-line solid phase extraction coupled to high performance liquid chromatography and tandem mass spectrometry” (NHANES, 2014). Detection limits in 100 μL of urine were as follows: BPA, 0.2 ng/mL; BPF, 0.2 ng/mL; BPS, 0.1 ng/mL.

## 2.4 Thyroid function measures

Self-reported recall data on thyroid function were recorded by an interviewer during the home visit by trained interviewers using the Computer-Assisted Personal Interviewing (CAPI) system [(CDC 2012, 2014, 2016)](https://www.cdc.gov/nchs/nhanes/index.htm). All adult study participants were asked to report whether or not they currently have a clinician-diagnosed thyroid problem. In order to be classified with a thyroid problem for this study, a participant must have answered yes to both of the following questions: “Has a doctor or other health professional ever told you that you had a thyroid problem?” and “Do you still have a thyroid problem?”

Participants for whom thyroid function data could not be ascertained, either due to lack of knowledge or unwillingness to share health care information, were excluded from this study.

## 2.5 Covariate Selection

The covariates selected for adjustment were informed by previous bisphenol exposure studies in NHANES that described relevant possible confounders based on sociodemographic and lifestyle characteristics of a representative sample (Calafat et al., 2008; La-kind and Naiman 2011; Nelson et al., 2012). They included gender, age, race, income, education, BMI, and smoking. We also included creatinine as a covariate to account for urine dilution (Barr et al., 2005) and included it in our models as a continuous variable.

## 2.6 Statistical Analysis

*2.6.1. Nonlinear analysis*

Unadjusted and adjusted multivariable logistic regression models were fit to estimate associations between urinary markers of chemical exposures and current medical diagnosis of thyroid problems. For adjusted analyses, the same covariate set was used for all models (see Section 2.4).

Exposure variables were assessed for nonlinear associations with log-odds of current thyroid problems. Unadjusted and adjusted generalized additive models (GAMs) were fit separately for each exposure. Nonlinearity was also separately assessed in other models found below (i.e. quantile regression and PCA models), in case the nonlinearity depended on the operationalization of the exposure-outcome relationship.

*2.6.2. Quantile Regression*

All data analyses were completed using R statistical software, Version 3.6.3 (R Core Team, 2020). Adjusted and unadjusted quantile regression (QR) models at the 50th (p50, median) and 75th percentiles (p75, 3rd quartile) were employed to evaluate the association between self reported thyroid function outcome and urinary BPA, BPF and BPS. The quantiles of exposures implemented for this analysis were driven by exposure distributions trends and with the purpose of identifying any potential presence of effect modification at the given quantiles.

In this population with right skewed environmental bisphenols exposure, the mean values would not provide information on the right tail of the distribution that can also capture abnormal indicative of high risk of exposure. Adjusted models were fit for sex, BMI, smoking, creatinine and age.

*2.6.3. Principal Component Analysis*

To test our third hypothesis, that exposure to multiple synthetic organic chemicals (mixtures) is associated with increased prevalence of self-reported thyroid dysfunction, principal component analysis was run on the exposure data. Principal component analysis, or PCA, is a dimension reduction technique that utilizes an orthogonal transformation to create new, uncorrelated variables from the data that maximize variance (Jollife and Cadima, 2016). In addition to BPA, BPS, and BPF, the data set for this method included eight additional chemicals commonly found in personal care products, listed in table X. Variables were centered and scaled prior to analysis.

The principal components identified by the analysis were then used as individual predictors in multivariate logistic regression with covariates selected for previous models (see Section 2.4). Separate models were run for the top X principal components. Since PCA requires complete data for all exposure and covariate variables, NAs were removed from the dataset.

References:

Hartle, J. C., Navas-Acien, A., & Lawrence, R. S. (2016). The consumption of canned food and beverages and urinary Bisphenol A concentrations in NHANES 2003-2008. Environmental research, 150, 375–382. <https://doi.org/10.1016/j.envres.2016.06.008>

Johnson CL, Dohrmann SM, Burt VL, Mohadjer LK (2014). National health and nutrition examination survey: sample design, 2011-2014. Vital Health Stat 2. 2014(162):1-33.

Jolliffe IT, Jorge Cadima J (2016). Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A. 374:20150202. http://doi.org/10.1098/rsta.2015.0202.

Hornung, R. W., & Reed, L. D. (1990). Estimation of average concentration in the presence of nondetectable values. *Applied occupational and environmental hygiene*, *5*(1), 46-51.

NHANES - National Health and Nutrition Examination Survey (2014). 2013-2014 Data Documentation, Codebook, and Frequencies. Personal Care and Consumer Product Chemicals and Metabolites (EPHPP\_H). Accessed at <https://wwwn.cdc.gov/Nchs/Nhanes/2013-2014/EPHPP_H.htm>.

NHANES - National Health and Nutrition Examination Survey (2016). 2015-2016 Data Documentation, Codebook, and Frequencies. Personal Care and Consumer Product Chemicals and Metabolites (EPHPP\_I). Accessed at <https://wwwn.cdc.gov/Nchs/Nhanes/2015-2016/EPHPP_I.htm>.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Zhou X, Ye X, Calafat AM. Automated on-line column-switching HPLC-MS/MS method for the quantification of triclocarban and its oxidative metabolites in human urine and serum. J Chromatogr B Analyt Technol Biomed Life Sci. 2012 Jan; 881-882:27-33