Introduction

Perfluoroalkyl and polyfluoroalkyl substances (PFAS) are large and complex molecules produced in industry that are widespread, persistent, and bioaccumulative in our environment and have been proposed to cause adverse health effects (NIH, 2022). PFAS have been used inustriuously as surfactants, lubricants, paints, polishes, food packaging and flame retarding compounds (NIH, 2022). Perfluorooctanoic acid (PFOA) and perfluorooctane sulfonate (PFOS) are two species of PFAS that were manufactured for the longest period of time and are found most frequently in the environment (NIH, 2022). These PFAS are no longer used in the US, though they have been replaced with alternative species of PFAS (NIH, 2022). Exposure to PFAS can occur when these chemicals contaminate air, water and soil (NIH, 2022). The most common routes of exposure for humans to PFAS are ingesting food or water contaminated with PFAS, using products containing PFAS, or breathing air polluted with PFAS (NIH, 2022). Children are at highest risk of exposure to PFAS, as they ingest more water and food per body weight and are likely to crawl on the floor and put objects in their mouths and therefore be exposed to PFAS in household dust, carpets, cleaning products and toys (EPA, 2022). The CDC's National Health and Nutrition Examination Survey (NHANES) in the year 2011-2012 found evidence of PFAS in the serum of 97% of participants (CDC, 2020).

PFAS occur as complex mixtures in the environment (CDC, 2020). Synthesis of PFAS between the 1950s and 2000s used electrochemical fluorination (ECF), which resulted in linear and branched isomers (NHANES, 2020). Fluorotelomerization, another method of PFAS synthesis, resulted in exclusively linear isomers (NHANES, 2020). The PFAS analytes tested in the NHANES assay are: perfluorodecanoate (PFDeA), perfluorohexane sulfonate (PFHxS), perfluorononanoate (PFNA), 2-(Nmethyl-perfluorooctane sulfonamido) acetate (PFOSA-AcOH), perfluoroundecanoate (PFUA), sum of branched perfluorooctanoate isomers (Sb-PFOA), and sum of perfluoromethylheptane sulfonate isomers (Sm-PFOS) (NHANES, 2020).

PFAS have been studied in association with several adverse health outcomes, including interference with thyroid function (Coperchini et al, 2021). It has been proposed that PFAS has adverse effects on the thyroid gland (Coperchini et al, 2021). Disruption of thyroid hormones could cause effects on the cardiovascular system, fertility, and fetal development (Coperchini et al, 2021). A meta-analysis of PFAS and thyroid function concluded that PFAS has a negative effect on total T4, and this effect varied by PFAS concentration (Kim et al, 2018). The association between PFAS and thyroid function remains inconclusive. The objective of our report is to analyze the PFAS analyte concentrations in blood samples collected and measured by NHANES in relation to participants' experience with thyroid problems.

Methods

Data Compilation

Data were compiled from the United States Center for Disease Control and Prevention's National Health and Nutrition Examination Survey (NHANES). Two cycles worth of data were

utilized for this analysis, 2017-2018 and 2015-2016. We used lab data from NHANES for PFAS concentration that was collected and measured in the participant's serum. Participants were 12 years or older and chosen by a ⅓ sample (n= 3200).

The concentrations of 9 PFAS's analytes were measured. Concentrations of linear PFOA (n-PFOA), sum of branched isomers of PFOA (Sb-PFOA, branched PFOA isomers), linear PFOS (n-PFOS), and sum of perfluoromethylheptane sulfonate isomers (Sm-PFOS, monomethyl branched PFOS isomers) were measured. These analytes were separated and measured by liquid chromatography and mass spectrometry

We used a binary measure for thyroid dysregulation. We used data from the question on NHANES, "Has a doctor or other health professional ever told you that you had a thyroid problem," to which participants could answer "yes" or "no." Thus, our definition of thyroid disease included conditions characterized by both hypothyroidism, an underfunctioning thyroid, and hyperthyroidism, an over-active thyroid.

Data Wrangling

Three datasets were downloaded from NHANES for each cycle, the lab dataset containing PFAS concentrations, a demographic dataset containing age information, and a medical questionnaire dataset containing the information on the outcome of interest. Once downloaded from the NHANES website, R programming language was utilized to compile these data into a workable dataset and for all subsequent statistical analysis.

Data Cleaning

Once in a complete working dataset, the laboratory PFAS concentrations were assessed for normal distribution. All laboratory data was significantly right-skewed. Subsequently, laboratory PFAS concentrations were log transformed to result in a more normal distribution. PFAS variables were also analyzed for their percent below the limit of detection (LOD). Three analytes had a significant proportion of samples which were below the LOD. These three were PFUA, Sb-PFOA, and PFOSA-AcOH, which were marked for exclusion in the majority of analyses (Table 1). Lastly, the LOD for NHANES laboratory data is coded at 0.07, however these values were re-coded as the LOD/square root of two (= 0.0495) for all nine variables.

Principal Component Analysis, Heatmap, and Correlation Plot:

Principal Component Analysis (PCA) was utilized to understand the interactions within the PFAS mixture. The "pls" package was used to conduct the PCA, by utilizing the prcomp function. A heatmap of participant identifier on the y-axis and PFAS concentration on the x-axis was developed to demonstrate the distribution of concentration levels across the study population. The "tidyverse" package was used to construct the heatmap, utilizing the ggplot and geom_tile functions. Lastly, to identify which PFAS analytes were correlated with each other, a correlation density plot was constructed using the "ggcorrplot" package.

Logistic Regression:

Logistic Regression was utilized to statistically assess the relationship between thyroid dysregulation and PC1 & PC2. The "glm" function was utilized from the Base R "stats" package.

The first analysis was conducted unadjusted, with no other variables included. The second analysis was conducted with age as an adjusted variable, in addition to the original variables.

Results

Figure 1 shows the percentage of variance explained by the PCA. PC1 and PC2 showed to be the most influential components. PC1 explained 68.1% of variance, and PC2 explained 14.49% of variance. *Figure 2* shows the effect of PC1 and PC2 by age group.

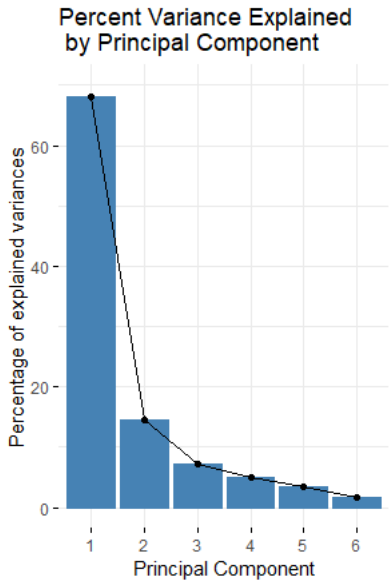

Figure 1. Scree plot showing percent variance explained by PCA.

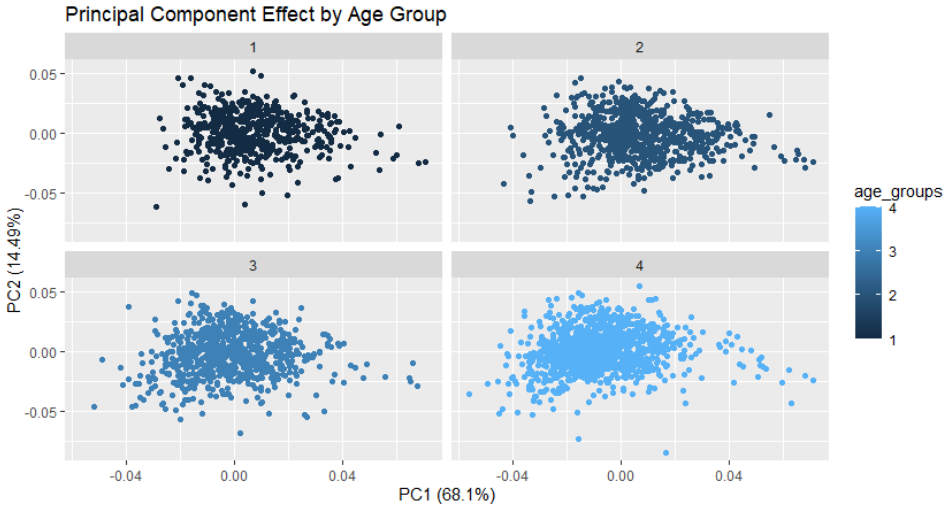

Figure 2. Scatter plot showing principal component effect by age.

Figure 3 shows the Principal Fitting Components (PFC) for the PC1 and PC2. PC1 is inversely associated with each of the component PFAS species. PC1 increases when the log concentrations of the analytes PFDeA, PFHxS, PFNA, n-PFOA, n-PFOS and sm-PFOS decrease. PC2 is positively associated with the log concentrations of PFHxS, n-PFOA, and sm-PFOS. PC2 is inversely associated with the log concentrations of PFDeA, PFNA, and n-PFOS. *Figure 4* shows a correlation plot of the PFCs, and *Figure 5* shows a heat map of the concentrations of the 9 PFAS species. There does not appear to be a significant correlation between the concentrations of separate PFAS analytes. Figure 6 shows the results of the logistic regression between PC1 and PC2 with thyroid disruption. As shown by Figure 6A, the relationship between PC1 and PC2 with thyroid disruption are not significant. However, as shown by Figure 6B, when we adjust for age the relationship between PC1 and thyroid disruption is shown to be significant ($p < 0.01$).

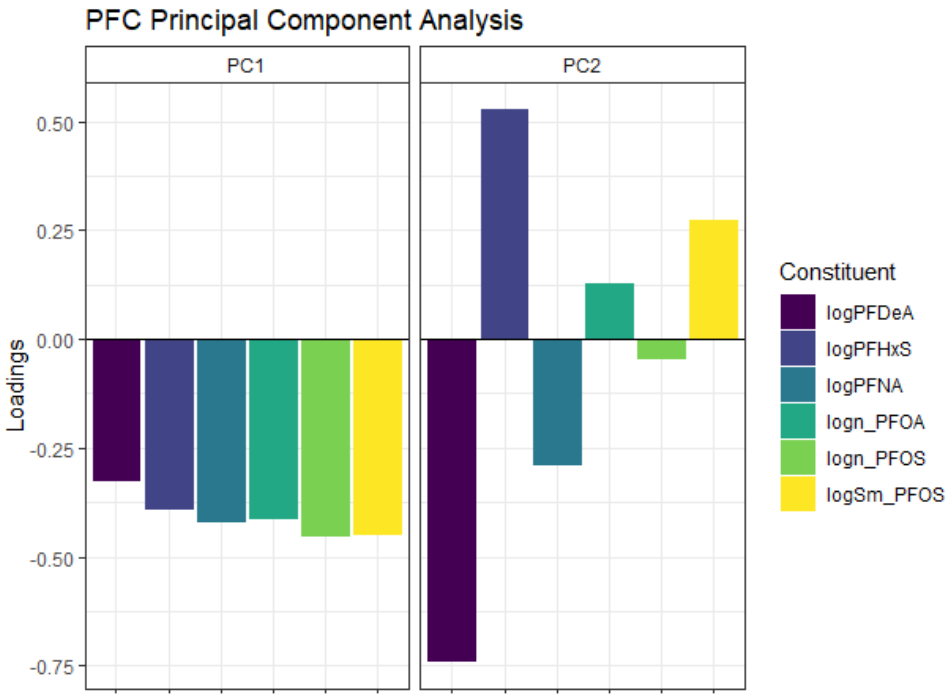

Figure 3. Loadings showing the PFC of PC1 and PC2.

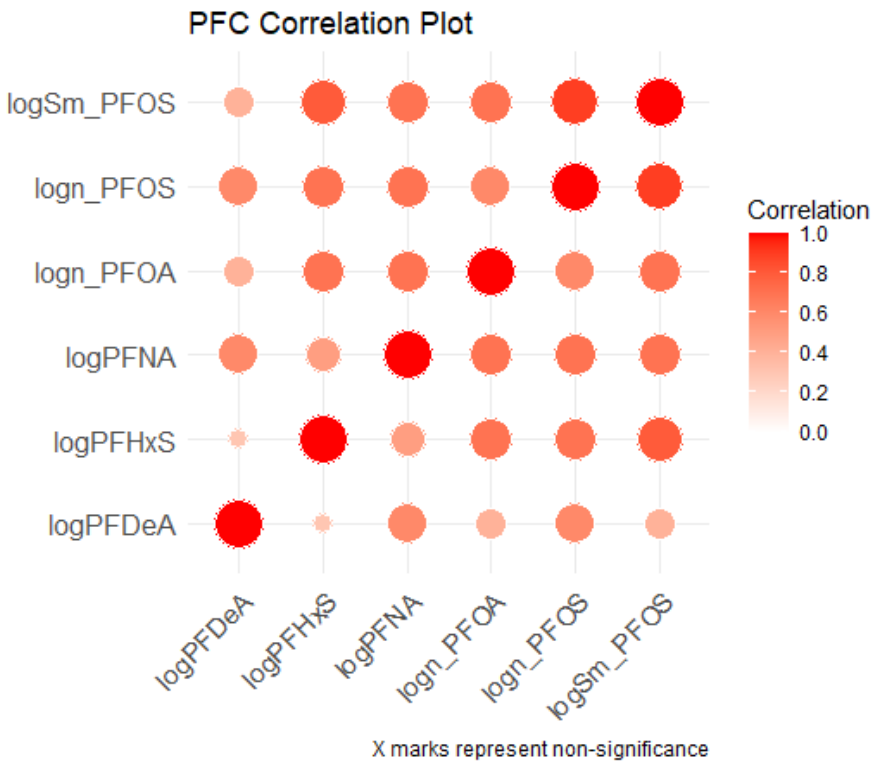

Figure 4. PFC Correlation Plot.

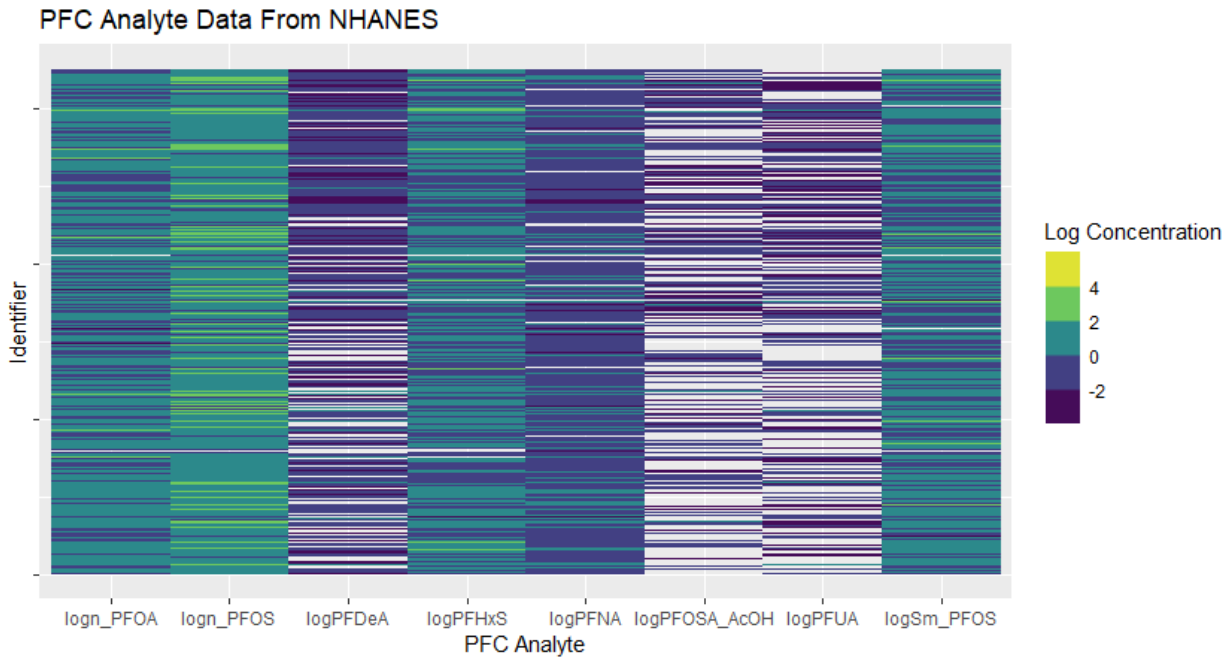

Figure 5. Heat map of PFC analytes data in log concentrations.

A.

Logistic Regression, unadjusted			
	Intercept	PC1	PC2
Coefficients	0.1162	0.001104	0.006748
P-values	<0.00001	0.69145	0.26299

B.

Logistic Regression, adjusting for age				
	Intercept	**PC1**	PC2	**Age**
Coefficients	-0.07511	**0.013108**	0.002162	**0.003813**
P-values	<0.00001	<0.00001	0.71523	<0.00001

Figure 6. Logistic regression for PC1 and PC2 (A) and the same logistic regression while adjusting for age (B).

Discussion

The results of the PCA showed varying results that were mostly insignificant. The PCA showed a large percentage of variance explained by the components PC1 and PC2. PC1 was interpretable, and was shown to increase when the log concentrations of the analytes PFDeA, PFHxS, PFNA, n-PFOA, n-PFOS and sm-PFOS decrease. PC2 is less interpretable, as PC2 is positively associated with the log concentrations of PFHxS, n-PFOA, and sm-PFOS, and inversely associated with the log concentrations of PFDeA, PFNA, and n-PFOS. The difference in directions of the loadings for PC2 may reflect that thyroid disruption can be caused by hyper or hypo regulation of hormones.

The correlation plot and heat map of the concentrations of the nine PFAS analytes showed that there was no apparent clustering or correlation between the different analystes, which allowed for our analysis to avoid any confounding relationships between PFAS mixture components. The logistic regression for PC1 and PC2 with thyroid dysfunction were not significant. However, when we adjusted for age, the relationship between PC1 and thyroid dysfunction was shown to be significant. Current understanding of PFAS exposure are consistent with this finding, as recent studies have hypothesized an effect on thyroid function caused specifically by early life exposure to PFAS. A systematic review found some evidence of a positive association between TSH and PFNA in boys below the age of eleven (Ballesteros et al, 2016).

There were several limitations to the analysis of PFAS and thyroid disruption. Firstly, the values for PFAS analyte concentrations from the lab data seemed to be rounded and were not precise, making our analysis less sensitive. Secondly, we only chose to use the measure of thyroid disruption as the question on NHANES “has a doctor or other health professional ever told you have thyroid disease” instead of exploring all possible downstream effects of thyroid hormone level changes. Disruption to the thyroid gland can cause effects including effects on the cardiovascular system and fertility (Coperchini et al, 2021). Notably, since thyroid diseases

can occur in a “positive” or “negative” direction, that is, be characterized by either overactivity or underactivity, a relationship between PFAS exposure and thyroid dysregulation could have been obscured by including both “directions” of thyroid disorder; we hypothesize that some PFAS species have a suppressive effect on thyroid function and some have an activating effect on the thyroid, or that PFAS mixtures a whole can be suppressive or activating, and that there might be an association with PFAS and either hypo- or hyperthyroidism groups if analyzed as separate groups.

Conclusion

The results of the PCA and g-computations do not show conclusive evidence connecting PFAS exposure and incidence of thyroid dysregulation. PC1 was interpretable and showed an inverse relationship between the log concentrations of analytes PFDeA, PFHxS, PFNA, n-PFOA, n-PFOS and sm-PFOS. PC2 was not interpretable, as three analytes showed a positive association with PC2 and three analytes showed an inverse relationship with PC2. Upon adjusting the logistic regression for age, the relationship between PC1 and thyroid dysregulation was shown to be significant. However, this relationship was contradictory to our original hypothesis, such that with increasing PFAS concentrations there was a decrease in thyroid dysregulation. It is possible that inclusion of the mixture of PFAS chemicals obscured a potential relationship, given that most analytes had low concentration levels. Thus, further research is needed to understand if particular PFAS chemicals outside of mixtures have effects on thyroid regulation.

References

1. U.S. Department of Health and Human Services. (n.d.). Perfluoroalkyl and Polyfluoroalkyl Substances (PFAS). National Institute of Environmental Health Sciences. Retrieved April 17, 2022, from https://www.niehs.nih.gov/health/topics/agents/pfc/index.cfm
2. Coperchini, F., Croce, L., Ricci, G., Magri, F., Rotondi, M., Imbriani, M., & Chiovato, L. (2021, January 19). Thyroid disrupting effects of old and new generation pfas. Frontiers in endocrinology. Retrieved April 16, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7851056/
3. Environmental Protection Agency. (n.d.). EPA. Retrieved April 17, 2022, from https://www.epa.gov/pfas/our-current-understanding-human-health-and-environmental-risks-pfas
4. Kim, M. J., Moon, S., Oh, B.-C., Jung, D., Ji, K., Choi, K., & Park, Y. J. (2018, May 10). Association between perfluoroalkyl substances exposure and thyroid function in adults: A meta-analysis. PloS one. Retrieved April 16, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5945046/
5. Centers for Disease Control and Prevention. (2020, November). National Health and Nutrition Examination Survey PFAS_J. Centers for Disease Control and Prevention. Retrieved April 16, 2022, from https://wwwn.cdc.gov/Nchs/Nhanes/2017-2018/PFAS_J.htm

6. Keil, A. P., Buckley, J. P., O'Brien, K. M., Ferguson, K. K., Zhao, S., White, A. J. 2020. A quantile-based g-computation approach to addressing the effects of exposure mixtures. *Environmental Health Perspectives* 128*(4)* https://doi.org/10.1289/EHP5838.
7. Ballesteros, V., Costa, O., Iñiguez, C., Fletcher, T., Ballester, F., & Lopez-Espinosa, M.-J. (2016, November 22). *Exposure to perfluoroalkyl substances and thyroid function in pregnant women and children: A systematic review of Epidemiologic Studies*. Environment International. Retrieved April 17, 2022, from https://www.sciencedirect.com/science/article/pii/S0160412016306195