The association between mixed exposure to per- and polyfluoroalkyl substances (PFAS) and the risk of diabetes

Abstract

Introduction
This study aims to investigate the relationship between exposure to per- and polyfluoroalkyl substances (PFAS) and the risk of diabetes by analyzing a large dataset from the National Health and Nutrition Examination Survey (NHANES).

Material and methods
This study analyzed data from the US NHANES database from 2003 to 2018. Generalized linear regression models were used to examine the relationship between serum chemical concentrations and diabetes. Weighted quantiles and regression were employed to assess the association between mixed chemical exposure and diabetes. The 'mixture effect' between chemicals and diabetes was estimated using a Quantile g-Computation model.

Results
The study involved 11,780 participants, with 1,451 individuals having diabetes and 10,329 being non-diabetic. Logistic regression analysis was to investigate the relationship between five PFAS chemicals and the diabetes risk, while controlling for all covariates. The results indicated that none PFAS chemicals exhibited a statistically significant association with an increased risk of diabetes. However, when exploring the positive and negative constraint models of the Weighted Quantile Sum (WQS), the coefficient of the chemical mixing index displayed a strong correlation with diabetes in model III, even after adjusting for covariates. Additionally, an evaluation of the 'mixture effect' using Quantile g-computation revealed that among the five chemicals, the combined exposure effect of Perfluorononanoate (PFNA) on diabetes was positive, whereas exposure to the other four chemicals had a negative impact.

Conclusions
The study findings suggest a possible connection between PFAS and an increased risk of diabetes, particularly highlighting PFNA's positive correlation with the combined exposure effect on diabetes.

The association between mixed exposure to per- and polyfluoroalkyl substances (PFAS) and the risk of diabetes

Abstract

Introduction: This study aims to investigate the relationship between exposure to per- and polyfluoroalkyl substances (PFAS) and the risk of diabetes by analyzing a large dataset from the National Health and Nutrition Examination Survey (NHANES).

Material and methods: This study analyzed participants with complete data from the US NHANES database spanning from 2003 to 2018. Generalized linear regression models were used to examine the relationship between serum chemical concentrations and diabetes. Weighted quantiles and regression were employed to assess the association between mixed chemical exposure and diabetes. The 'mixture effect' between chemicals and diabetes was estimated using a Quantile g-Computation model.

Results: The study involved a total of 11,780 participants, with 1,451 individuals having diabetes and 10,329 being non-diabetic. Logistic regression analysis was utilized to investigate the relationship between five PFAS chemicals and the risk of diabetes, while controlling for all covariates. The results indicated that none of the PFAS chemicals exhibited a statistically significant association with an increased risk of diabetes. However, when exploring the positive and negative constraint models of the Weighted Quantile Sum (WQS), the coefficient of the chemical mixing index displayed a strong correlation with diabetes in model III, even after adjusting for covariates. Additionally, an evaluation of the 'mixture effect' using Quantile g-computation revealed that among the five chemicals, the combined exposure effect of Perfluorononanoate (PFNA) on diabetes was positive, whereas exposure to the other four chemicals had a negative impact.

Conclusions: The study findings suggest a possible connection between PFAS and an increased risk of diabetes, particularly highlighting PFNA's positive correlation with the combined exposure effect on diabetes.

Keywords PFAS; Mixed exposure; Diabetes; Weighted quantiles sum (WQS)

Introduction

Diabetes, a chronic disease, affects around 537 million adults globally and is expected to exceed 600 million by 2030, with type 2 diabetes accounting for about 90% of cases [1]. Diabetes mellitus (DM) comprises a group of chronic metabolic disorders characterized by high blood glucose levels, leading to various complications over time. Untreated DM can cause damage to blood vessels, nerves, tissues, and organs[2]. Epidemiological studies have highlighted the significant roles of genetics and lifestyle in the development of type 2 DM[3-5]. Factors such as physical inactivity, poor diet, smoking, and high BMI contribute to the rising prevalence of diabetes. Additionally, emerging research suggests that environmental pollutants, such as perfluoroalkyl and polyfluoroalkyl substances, along with other endocrine-disrupting compounds, may impact the progression of diabetes [6].

PFASs are extensively utilized in a wide range of commercial applications, including surfactants, lubricants, and flame retardant foams,[7] due to their exceptional stability and hydrophobic and oleophobic properties. Some PFASs are also employed as polymers in various industrial and consumer products, like waterproof coatings on textiles and non-stick coatings on kitchenwar However, PFASs are recognized as Persistent Organic Pollutants (POPs) because of their prolonged environmental presence and resistance to degradation [9-11]. They have been found in water sources, soil, plants, animals, and humans. Human exposure to PFASs can occur through different pathways, with the main one being contact with contaminated drinking water or food[11,12]. Several studies have indicated a potential link between PFAS chemicals, such as perfluorooctanoic acid (PFOA), PFOS, and others, and various adverse health effects, including preeclampsia, altered aminotransferase levels, elevated blood lipids, reduced antibody responses to vaccines, and low birth weight. However, the causality of these associations remains to be established [13].Gui et al. conducted a systematic review and meta-analysis of epidemiological evidence, revealing that PFAS exposure

is associated with an increased risk of type 2 diabetes mellitus. They specifically noted a parabolic dose-response relationship with PFOA exposure[14]. While the association between polychlorinated biphenyls (PCBs) and type 2 diabetes risk has been extensively studied over the years, the levels of PCBs in the environment and human body are gradually decreasing. As a result, there have been limited investigations into the relationship between PFASs and diabetes. Therefore, analyzing the relationship between mixed exposure to PFASs and diabetes is of significant importance.

Our study utilized data from the NHANES spanning from 2003 to 2018. To examine the association between five PFAS concentrations and diabetes, we employed a weighted multivariate logistic regression model. Prior to analysis, both ln transformation and quartile transformation were applied. Additionally, we utilized the WQS regression model to assess the relationship between mixed PFAS exposure and diabetes, allowing us to identify the specific PFAS that were more influential in the context of diabetes. To estimate the 'mixture effect' relationship between PFAS exposure and diabetes, we employed the Quantile g-Computation model. Our investigation aimed to contribute new epidemiological evidence to the field by exploring the relationship between mixed PFAS exposure and the risk of diabetes.

Materials and Methods

Study Population

The NHANES is a cross-sectional study conducted in the United States since the 1960s to assess the health and nutritional status of individuals, including children and adults. Approval was obtained from the Ethical Review Board of the National Center for Health Statistics, and informed consent was secured from all participants. Data from publicly available sources, spanning from 2003 to 2018, were collected for this study. A total of 80,312 participants were initially included, with 63,771 excluded due to missing PFAS level information. Of the

remaining 16,541 participants with PFAS measurements, 4,761 were further excluded for missing covariates. Fortunately, no outcome variables were missing. Ultimately, 11,780 participants were included in the final analysis, all of whom had complete information on both PFAS measurements and covariates. Further details on participant identification can be found in Figure 1.

Definition of Diabetes

Diabetes diagnoses were made by medical professionals, and information on diabetes status was collected through a questionnaire. Participants aged 20 years or older were asked if they had ever been diagnosed with diabetes by a doctor or healthcare professional. Positive responses to this question were used to identify individuals with diabetes in the survey.

Pollutant measurement

PFASs have become extensively utilized and have been found in various environmental sources, including water and organisms [15-17]. Humans can come into contact with PFASs through inhalation, skin exposure, and consuming contaminated food. Once in the body, PFASs initially circulate in the bloodstream and undergo metabolic changes, leading to the production of metabolites that can accumulate in specific organs and trigger toxic responses, such as endocrine disruption[18,19]. In this research, blood samples from participants were analyzed for five specific PFASs: Perfluorodecanoate (PFDeA), Perfluorohexane sulfonate (PFHxS), 2-(N-methyl-perfluorooctane) sulfonamido acetate (Me-PFOSA-AcOH), Perfluorononanoate (PFNA), and Perfluoroundecanoate (PFUA). These PFASs were identified as persisting in the environment from 2002 to 2018 using high-performance liquid chromatography coupled with tandem mass spectrometry (MS/MS). When the limit of detection (LOD) for a chemical was below 50%, the LOD logarithm/2 method was used to calculate the corresponding data, as 86% of the chemicals met this criterion.To ensure quality, procedural blanks and spiked

samples were included in each batch of samples to monitor potential contamination and analytical performance. Detection limits were set for each PFAS analyte, and measurements below these limits were processed using the NHANES LOD/2 method, a standard practice for handling non-detectable values in environmental exposure assessment.

Covariates

A comprehensive review of the relevant literature was conducted to identify potential covariates associated with exposure to Perfluoroalkyl and PFAS and the risk of diabetes. The selected covariates were categorized as either continuous or categorical variables. Continuous variables included age, family poverty index ratio (PIR), annual household income, and BMI. Categorical variables encompassed gender (male, female), race (Mexican American, other Hispanic, non-Hispanic white, non-Hispanic black, other race - including multiracial), education (less than ninth grade, grades 9-11 (including grade 12 without a diploma), high school /GED or equivalent, college or AA degree, associate's degree or above, denial, don't know), and marital status (married, widowed, divorced, separated, never married, living with a partner).

Statistical Methods

Descriptive statistics were used to summarize the demographic characteristics of the participants and concentrations of biomarkers. Urine creatinine was not needed as a correction factor since it is a laboratory measure of blood and does not require urine dilution. Categorical variable data were presented as the number of cases (percentage) and underwent testing. Non-normal continuous variable data were presented as the median (interquartile range) [M (Q1, Q3)], and analyzed using the Wilcoxon rank sum test. Normal continuous variable data were presented as mean $\pm$ standard deviation, and analyzed using the independent sample t-test. A natural log transformation (ln transformation) was applied to PFAS to establish a normal distribution. Spearman's

6

correlation was used to assess the relationships between chemicals.

A weighted multivariate logistic regression model was utilized to investigate the correlation between natural logarithm (ln) transformed and quartile transformed chemical concentrations and the incidence of diabetes. In order to assess the relationship between combined chemical exposure and diabetes, as well as to pinpoint the key chemicals influencing the occurrence of diabetes, a WQS regression model was employed. The WQS model is a statistical method that assesses the impact of each component of environmental exposure on the overall effect using a weighted index. This model has been widely used in evaluating the effects of exposure to mixtures[20]. The WQS model calculates quantiles for chemicals and assigns individual weights based on their relative importance within the mixture, facilitating the identification of potentially harmful substances.

The study utilized the Quantile g-Computation model to investigate the association between chemicals and diabetes by evaluating the 'mixture effect.' Quantile G-computation is a unique method for analyzing combined environmental exposures, as it estimates the parameters of marginal structural models and offers causal effect estimates. This model predicts the anticipated change in potential outcomes if a simultaneous intervention is implemented for all exposures, potentially while considering confounding factors. Accurate estimates of the true effect can be achieved through quantile g calculations when the assumptions of exchangeability, causal consistency, positivity, no interference, and correct model specification are met. Statistical analysis was performed using R3.4.3, with a two-sided P-value below 0.05 considered statistically significant [21].

Note: Due to the high proportion of samples with LOD ≤ 50%, they were not excluded from the study population. Instead, these samples were adjusted by dividing them by 2, and then missing values of independent variables were excluded. The NHANES protocol was approved by the NCHS research ethics review board, and all participants provided informed consent.

Results

Baseline characteristics of the study population

A total of 11,780 participants were included in this study for statistical analysis. The general characteristics of the study population can be found in Table 1. Among these participants, 1,452 were diagnosed with diabetes, representing a prevalence of 12.32%. The remaining 10,329 participants were diagnosed as non-diabetic, accounting for 87.68% of the sample. Statistical analysis indicated significant differences in age, race, education level, marital status, family income, BMI, and diabetes status between patients with and without diabetes ($P <$ 0.05). However, there was no significant difference in gender between the two groups ($P > 0.05$).

Distribution of chemical exposure and its association

The nuclear density distribution was analyzed for all five chemicals, and the results are illustrated in Figure 1 in the appendix. Figure 1 in the appendix shows that PFUA displays three distinct peaks, while PFNA, MPAH, PFHxS, and PFAeA each show a single prominent peak.

A Spearman correlation analysis was conducted to investigate the relationship between the measured values of the five chemicals, as shown in Figure 2. The results indicated that the correlation coefficient between PFUA and PFDeA was the highest at 0.76, suggesting a strong positive correlation. Furthermore, the correlation coefficient between PFUA and MPAH was 0.69, and between PFDeA and MPAH was 0.66, both showing significant positive correlations. On the other hand, the correlation coefficients for the remaining chemicals were below 0.2, indicating either a very weak correlation or no correlation at all. This hints at the possibility of high collinearity among the variables, necessitating the use of multivariate logistic regression to evaluate and confirm the collinearity.

Weighted generalized linear regression

Univariate logistic analysis was conducted using three models, each adjusted for specific covariates. The results are available in Annex Tables 1, 2, and 3. Model 1 included adjusted covariates of gender and age. Model 2 expanded on this by including additional covariates such as race, PIR, education level, and BMI. Lastly, Model 3 incorporated sex, age, race, PIR, education level, BMI, marital status, and annual household income as adjusted covariates.

None of the five chemicals showed a significant odds ratio (OR) for diabetes in all three models (P > 0.05). Specifically, in the adjusted model 3, the highest quartile (Q4) of PFDeA, PFHxS, and PFNA had values above 1, but none were statistically significant (P > 0.05). However, it is important to note that the lack of significance could be due to the small sample size and imbalance in the distribution of disease and non-disease populations in the study. Therefore, the lack of significance in this study should be interpreted with caution, as it may be influenced by these factors.

In the multivariate weighted logistic regression models, all five chemicals showed a variance inflation factor (VIF) of less than 10, suggesting no multicollinearity among them (refer to Table 2). The OR of PFNA on the outcome was calculated to be 1.14 (0.91, 1.43) in the multivariate weighted logistic regression model. Nevertheless, this relationship was not deemed statistically significant (P = 0.25).

Weighted quantiles and regression (WQS)

In the positive constraint model of the WQS, the coefficient for the chemical mixture index showed a weak association with diabetes (OR=1.02), which was not statistically significant. However, strong correlations between the chemical mixture index coefficient and diabetes were observed in Model 1, Model 2, and Model 3 after adjusting for relevant covariates (Model 1: OR=1.77, Model 2: OR=1.83, Model 3: OR=1.83). The results are presented in Table 4. Similarly, in the unadjusted model of the negative constraint in the WQS, the OR was 1.98,

but it was not statistically significant (P > 0.05). However, after adjusting for covariates, a robust association between the chemical mixture index and diabetes was evident in Model 1, Model 2, and Model 3 (Model 1: OR=1.87, Model 2: OR=1.90, Model 3: OR=1.90). These findings are summarized in Annex Table 5.

The first model (Model 1) considered two covariates: gender and age. The second model (Model 2) included a wider range of covariates, such as gender, age, race, PIR, education level, and BMI. Finally, the third model (Model 3) integrated additional covariates, including sex, age, race, PIR, education, BMI, marital status, and annual household income.

In the positive constraint model of WQS, the substances MPAH, PFHxS, and PFUA are considered to be relatively significant, with weights exceeding 1/18 each. Similarly, within the negative constraint model of WQS, both PFDeA and PFNA are deemed to be relatively important, with weights surpassing 1/18 for each chemical. For more detailed information, please consult Figure 3.

Quantile g-computation "Mixture effect" evaluation

For further information on the use of binary results in Q g-comp using the Q g-comp.no boot function, please refer to Schedule 6. The Q g-comp.boot function produces a marginal OR. It is important to understand that introducing nonexposed covariates, such as confounders, into the model will result in estimations of these parameters that are not equivalent. This is because the OR is non-collapsible. Marginal parameters offer estimates of population-average exposure effects, which are typically more interpretable and meaningful than conditional odds.

Further adjustments to the covariates were made using Model I, Model II, and Model III. The results of these adjustments are illustrated in Figure 8, showcasing four different models (1: no adjustment; 2: Model one; 3: Model two; 4: Model three) and their impact on the overall effect for each exposure. It is important to highlight that the constraints imposed on the weights in the WQS model may lead to biased estimations of the effects. In

contrast, the Q g-comp model allows for weights to shift in different directions, indicating the potential for some exposures to be beneficial while others may be harmful. This flexibility also addresses sampling variation that can occur with small or medium-sized samples. Within the Q g-comp model, the weights represent the proportion of effects aligning in the same direction, although in certain instances, they may correspond to a smaller or larger proportion compared to the overall 'mixed' effect. Notably, Figure 4 illustrates that the combined exposure effect of PFNA on diabetes is positive in all four models, while exposure to the other four chemicals shows negative effects.

Nonlinearities in the data were addressed using Qg-comp and Qg-comp. BOOT methods were applied to develop a model that included an interaction term and quadratic terms for each predictor of chemical exposure. By simultaneously accounting for the nonlinear effects of all exposures, a distinct nonlinear trend in the overall exposure effect was observed, as shown in Supplementary Figures 2 and 3. Importantly, the smoothed regression line aligns with the confidence interval of the marginal linear model, indicating consistency. Additionally, it is clear that the OR for the prevalence of PFAS and diabetes increases steadily with each quantile change.

Discussion

This study utilized a large sample of data to investigate the potential link between mixed exposure of serum perfluoroalkyl and PFAS and diabetes. The statistical analysis of 11,780 participants, revealing a diabetes prevalence of 12.32%, uncovered patterns in the baseline characteristics of the study population. The analysis brought to light a strong positive correlation between PFUA and PFDeA, indicating a potential high collinearity between these PFAS chemicals. This finding underscores the need for further investigative research that could have significant implications for environmental health policies and understanding of endocrine disorders. Additionally, the Weighted Quantile Sum regression analysis, after adjusting for covariates, demonstrated a robust association between the chemical mixture index and diabetes in all models. This

suggests a complex, potentially non-linear relationship between PFAS exposure and diabetes, offering new insights that could guide future epidemiological studies and risk assessment models. The findings indicate that PFAS exposure may be associated with an elevated risk of developing diabetes, particularly showing a positive correlation between mixed exposure of PFNA.

The association between PFAS exposure and diabetes is still a topic of debate in the research community, with no consensus reached yet[22]. However, several studies have shown a positive correlation between serum PFAS levels and elevated blood glucose indicators in individuals with type 2 diabetes[23,24]. For example, Cardenas et al. analyzed participants in the Diabetes Prevention Program (DPP) from 1996 to 1999 and found that high PFOS and PFOA levels were linked to increased insulin resistance (HOMA-IR) at baseline. They also found associations between high PFOS and PFOA levels and elevated HOMA-β, fasting proinsulin, and glycated hemoglobin A1c (HbA1c) levels at baseline. Yet, the follow-up survey revealed a weak correlation between PFAS exposure and these blood glucose indicators [24]. In a separate study in Tianjin, China, involving individuals aged 19-87 years, it was discovered that a 1% increase in serum PFOA and PFNA concentrations was significantly associated with higher fasting plasma glucose levels, while a 1% increase in PFAS, PFHxA, and PFHxS was linked to higher HbA1c levels [25].

Some studies have suggested a potential positive link between serum PFAS levels and an increase in diabetes measures, while others have found no significant or inverse associations. To delve deeper into this correlation, Liu et al. conducted an analysis using NHANES data from 2013 to 2014. Their findings indicated that branched-chain PFOS and linear PFOA were significantly linked to a reduction in fasting plasma glucose. However, there was no notable association observed between PFAS and 2-hour plasma glucose (GTT), insulin levels, or HOMA-IR [26]. In a similar study, Nelson et al. analyzed NHANES data from 2003 to 2004 to investigate the relationship between PFAS (including PFOA, PFNA, PFOS, and PFHxS) and insulin resistance. Their analysis, however, did

not identify any significant association [27].

Numerous studies have explored the link between PFAS and diabetes occurrence, with results varying based on the specific PFAS compounds investigated. Lundin et al. analyzed the diabetes prevalence in individuals with high PFOA exposures between 1947 and 1997. Their findings indicated a heightened risk of diabetes and diabetes-related mortality in those with moderate exposures, while no notable correlation was observed at low or high exposures[28]. Conversely, some other studies have documented either no significant connection or a reverse association between PFAS exposure and type 2 diabetes risk[29-31].

Recent investigations into the pathophysiology of chemically-induced diabetes, particularly related to PFAS exposure, have revealed potential mechanisms. Studies indicate that PFAS could disrupt endocrine function, specifically insulin signaling pathways, leading to insulin resistance, a precursor to type 2 diabetes [32]. PFAS tend to accumulate in the liver, a crucial organ in glucose regulation, where they may disrupt glucose metabolism and insulin sensitivity[33]. Additionally, PFAS exposure has been associated with changes in lipid metabolism, which is closely linked to insulin resistance. The precise molecular pathways through which PFAS cause these effects are still being studied but could involve the modulation of nuclear receptors and interference with hormone secretion and function [34]. This disruption of metabolic and endocrine pathways may contribute to the development of diabetes, emphasizing the importance of further research into the specific biological mechanisms underlying PFAS-induced metabolic dysregulation.

The study benefits from using NHANES data in the United States, which offers a sizable and representative sample, and from serologically detecting five PFAS chemicals for subsequent statistical analysis, ensuring result reliability. However, there are limitations. Firstly, the cross-sectional nature of the NHANES data used in this study hinders establishing temporal or causal associations between mixed PFAS exposure and diabetes risk. Secondly, diabetes pathogenesis is intricate, with potential confounding

factors like genetics, diet, and lifestyle not considered in the analysis. Additionally, PFAS chemicals undergo adsorption-partition reactions with environmental media upon entry into the environment, leading to the generation of various metabolites upon entering the human body. Some of these metabolites interact with biological macromolecules, causing toxic effects like growth, reproductive, and endocrine toxicity. Despite this, the study did not investigate the potential impact of PFAS metabolites on diabetes prevalence. Hence, future research should delve into the link between chemical metabolites and diabetes risk, as well as explore the physiological mechanisms connecting mixed PFAS exposure and diabetes risk.

Conclusions

The results showed no evidence of multicollinearity among the five PFAS chemicals studied. It is important to highlight that PFAS compounds have been proposed to potentially increase the risk of diabetes. Furthermore, a positive correlation was found between PFNA and the combined exposure effect on diabetes.

Author contribution

Mei Yu: conceptualization, methodology, validation, formal analysis, funding acquisition, and writing-original draft preparation and supervision. Xuxia Chu: conceptualization, funding acquisition, and writing-review and editing and project administration. Yefei Fang: writing-review and editing, data curation, and visualization. Xiaoqing Zhan: visualization, supervision, and project administration. All authors read and approved the final manuscript.

Funding

Data availability

The data generated in this study are publicly available in National Health and Nutrition Examination Survey (NHANES) at https://www.cdc.gov/nchs/nhanes/about_nhanes.htm.

Ethical approval

The studies involving human participants were received and approved by National Health and Nutrition Examination Survey. The database is public, it is unified use of ethics approval and does not have a corresponding ethical number.

Consent to participate

The patients/participants provided their written informed consent to participate in this study.

Consent for publication

Not applicable.

Conflict of interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1.   Magliano, D.J. and E.J. Boyko, IDF diabetes atlas. 2022.

2.   Brownlee, M., Biochemistry and molecular cell biology of diabetic complications. Nature, 2001. 414(6865): p. 813-20.

3.   Kolb, H. and S. Martin, Environmental/lifestyle factors in the pathogenesis and prevention of type 2 diabetes. BMC Med, 2017. 15(1): p. 131.

4.   Sladek, R., et al., A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature, 2007. 445(7130): p. 881-5.

5.   Zeggini, E., et al., Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science, 2007. 316(5829): p. 1336-41.

6.   Qi, W., et al., Per- and Polyfluoroalkyl Substances and Obesity, Type 2 Diabetes and Non-alcoholic Fatty Liver Disease: A Review of Epidemiologic Findings. Toxicol Environ Chem, 2020. 102(1-4): p. 1-36.

7.   Buck, R.C., et al., Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins. Integr Environ Assess Manag, 2011. 7(4): p. 513-41.

8.   Buck, R.C., et al., Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins. Integrated environmental assessment and management, 2011. 7(4): p. 513-541.

9.   Zhang, Z., et al., Biodegradation of per- and polyfluoroalkyl substances (PFAS): A review. Bioresour Technol, 2022. 344(Pt B): p. 126223.

10.   Brusseau, M.L., R.H. Anderson, and B. Guo, PFAS concentrations in soils: Background levels versus contaminated sites. Sci Total Environ, 2020. 740: p. 140017.

11.   Domingo, J.L. and M. Nadal, Human exposure to per- and polyfluoroalkyl substances (PFAS) through drinking water: A review of the recent scientific literature. Environ Res, 2019. 177: p. 108648.

12. Clara, M., et al., Emissions of perfluorinated alkylated substances (PFAS) from point sources--identification of relevant branches. Water Sci Technol, 2008. 58(1): p. 59-66.

13. Magliano, D.J., E.J. Boyko, and I.D.F.D.A.t.e.s. committee, IDF Diabetes Atlas, in Idf diabetes atlas. 2021, International Diabetes Federation

© International Diabetes Federation, 2021.: Brussels.

14. Gui, S.-Y., et al., Association between per-and polyfluoroalkyl substances exposure and risk of diabetes: a systematic review and meta-analysis. Journal of exposure science & environmental epidemiology, 2023. 33(1): p. 40-55.

15. Nguyen, M.A., et al., Spatial distribution and source tracing of per- and polyfluoroalkyl substances (PFASs) in surface water in Northern Europe. Environ Pollut, 2017. 220(Pt B): p. 1438-1446.

16. Prevedouros, K., et al., Sources, fate and transport of perfluorocarboxylates. Environ Sci Technol, 2006. 40(1): p. 32-44.

17. Xie, L.N., et al., Concentration, spatial distribution, and health risk assessment of PFASs in serum of teenagers, tap water and soil near a Chinese fluorochemical industrial plant. Environ Int, 2021. 146: p. 106166.

18. Zheng, X.-q., et al., Growth inhibition and DNA damage in the earthworm (Eisenia fetida) exposed to perfluorooctane sulphonate and perfluorooctanoic acid. Chemistry and Ecology, 2016. 32(2): p. 103-116.

19. Caron-Beaudoin, É., et al., Exposure to perfluoroalkyl substances (PFAS) and associations with thyroid parameters in First Nation children and youth from Quebec. Environ Int, 2019. 128: p. 13-23.

20. Carrico, C., et al., Characterization of Weighted Quantile Sum Regression for Highly Correlated Data in a Risk Analysis Setting. J Agric Biol Environ Stat, 2015. 20(1): p. 100-120.

21. Schmidt, S., Quantile g-Computation: A New Method for Analyzing Mixtures of Environmental Exposures. Environ Health Perspect, 2020. 128(10): p. 104004.

22. Domazet, S.L., et al., Longitudinal Associations of Exposure to Perfluoroalkylated Substances in Childhood and Adolescence and Indicators of Adiposity and Glucose Metabolism 6 and 12 Years Later: The European Youth Heart Study. Diabetes Care, 2016. 39(10): p. 1745-51.

23. Kim, J.H., et al., The modifying effect of vitamin C on the association between perfluorinated compounds and insulin resistance in the Korean elderly: a double-blind, randomized, placebo-controlled crossover trial. Eur J Nutr, 2016. 55(3): p. 1011-20.

24. Cardenas, A., et al., Plasma Concentrations of Per- and Polyfluoroalkyl Substances at Baseline and Associations with Glycemic Indicators and Diabetes Incidence among High-Risk Adults in the Diabetes Prevention Program Trial. Environ Health Perspect, 2017. 125(10): p. 107001.

25. Duan, Y., et al., Distribution of novel and legacy per-/polyfluoroalkyl substances in serum and its associations with two glycemic biomarkers among Chinese adult men and women with normal blood glucose levels. Environ Int, 2020. 134: p. 105295.

26. Liu, H.S., et al., Association among total serum isomers of perfluorinated chemicals, glucose homeostasis, lipid profiles, serum protein and metabolic syndrome in adults: NHANES, 2013-2014. Environ Pollut, 2018. 232: p. 73-79.

27. Nelson, J.W., E.E. Hatch, and T.F. Webster, Exposure to polyfluoroalkyl chemicals and cholesterol, body weight, and insulin resistance in the general U.S. population. Environ Health Perspect, 2010. 118(2): p. 197-202.

28. Lundin, J.I., et al., Ammonium perfluorooctanoate production and occupational mortality. Epidemiology, 2009. 20(6): p. 921-8.

29. Donat-Vargas, C., et al., Perfluoroalkyl substances and risk of type II diabetes: A prospective nested case-control study. Environ Int, 2019. 123: p. 390-398.

30. MacNeil, J., et al., A cross-sectional analysis of type II diabetes in a community with exposure to

perfluorooctanoic acid (PFOA). Environ Res, 2009. 109(8): p. 997-1003.

31.   Conway, B., K.E. Innes, and D. Long, Perfluoroalkyl substances and beta cell deficient diabetes. J Diabetes

Complications, 2016. 30(6): p. 993-8.

32.   Zhang, L., et al., Perfluorooctane sulfonate acute exposure stimulates insulin secretion via GPR40 pathway.

Science of the total environment, 2020. 726: p. 138498.

33.   Sen, P., et al., Exposure to environmental contaminants is associated with altered hepatic lipid metabolism in

non-alcoholic fatty liver disease. Journal of hepatology, 2022. 76(2): p. 283-293.

34.   DeWitt, J.C., Toxicological effects of perfluoroalkyl and polyfluoroalkyl substances. 2015, Springer.

Figure legend

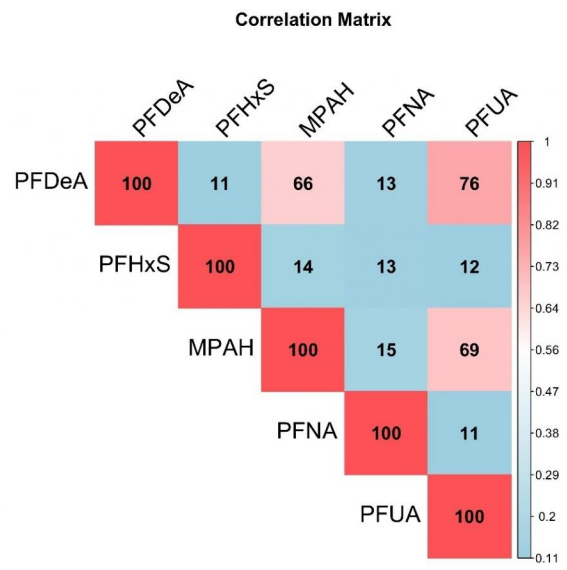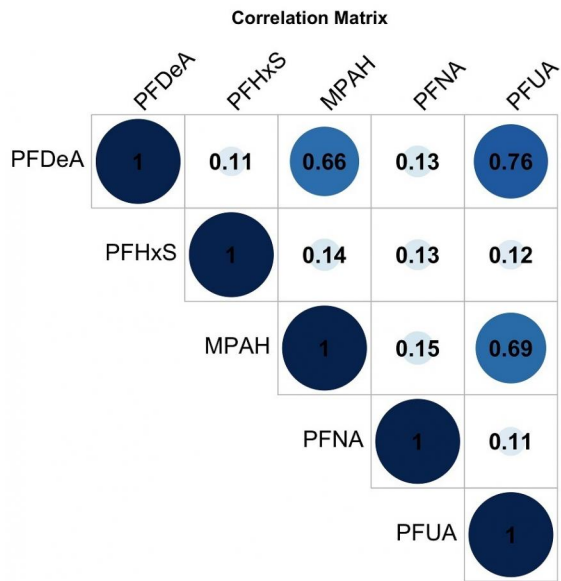**Fig. 1 Inclusion and exclusion path diagram**

Fig. 2 Spearman correlation coefficient between pairwise chemical measurements

Fig. 3 Weight of regression indicators on diabetes in the WQS model. (a) Forward constraints; (b) Negative constraints

Fig. 4 Q G-computation weights for the four models. (a) Model I; (b) Model II; (c) Model III; (d) Model 4

Correlation Matrix

	PFDeA	PFHxS	MPAH	PFNA	PFUA
PFDeA	1	0.11	0.66	0.13	0.76
PFHxS		1	0.14	0.13	0.12
MPAH			1	0.15	0.69
PFNA				1	0.11
PFUA					1

Correlation Matrix

|	PFDeA	PFHxS	MPAH	PFNA	PFUA
PFDeA	100	11	66	13	76
PFHxS		100	14	13	12
MPAH			100	15	69
PFNA				100	11
PFUA					100

Table 1 Population characteristics of diabetes in adults from the US National Health and Nutrition Examination Survey 2002-2018

	level	Non-Diabetes N=10329（87.68%）	Diabetes N=1451（12.32%）	P
Gender (%)	Male	4956 (48.0)	734 (50.6)	0.067
	Female	5373 (52.0)	717 (49.4)	
Age (median [IQR])		46.00 [32.00, 62.00]	63.00 [54.00, 71.00]	<0.001
Race (%)	Mexican American	1581 (15.3)	272 (18.7)	<0.001
	Other Hispanic	852 (8.2)	140 (9.6)	
	Non-Hispanic White	4756 (46.0)	527 (36.3)	
	Non-Hispanic Black	2083 (20.2)	370 (25.5)	
	Other Race	1057 (10.2)	142 (9.8)	
Family PIR (median [IQR])		2.18 [1.15, 4.13]	1.79 [1.03, 3.34]	<0.001
Education (%)	Less Than 9th Grade	965 (9.3)	282 (19.4)	<0.001
	9-11th Grade (Includes 12th grade with no diploma)	1398 (13.5)	266 (18.3)	
	High School Grad/GED or Equivalent	2372 (23.0)	314 (21.6)	
	Some College or AA degree	3128 (30.3)	372 (25.6)	
	College Graduate or above	2457 (23.8)	214 (14.7)	
	Refuse	3 (0.0)	0 (0.0)	
	Don't know	6 (0.1)	3 (0.2)	
BMI (median [IQR])		27.60 [24.00, 31.96]	31.05 [27.10, 36.30]	<0.001
Marital (%)	Married	5375 (52.0)	808 (55.7)	<0.001
	Widowed	741 (7.2)	222 (15.3)	
	Divorced	1049 (10.2)	201 (13.9)	

	Separated	329 (3.2)	55 (3.8)	
	Never married	1963 (19.0)	115 (7.9)	
	Living with partner	869 (8.4)	49 (3.4)	
	Refuse	3 (0.0)	1 (0.1)	
Annual household income (median [IQR])		7.00 [5.00, 11.00]	6.00 [4.00, 9.00]	<0.001
NHANES Cycles (%)	2003-2004	1191 (11.5)	145 (10.0)	<0.001
	2005-2006	1263 (12.2)	130 (9.0)	
	2007-2008	1369 (13.3)	190 (13.1)	
	2009-2010	1492 (14.4)	177 (12.2)	
	2011-2012	1218 (11.8)	171 (11.8)	
	2013-2014	1369 (13.3)	223 (15.4)	
	2015-2016	1258 (12.2)	203 (14.0)	
	2017-2018	1169 (11.3)	212 (14.6)	
Diabetes (%)	0	10329 (100.0)	0 (0.0)	<0.001
	1	0 (0.0)	1451 (100.0)	

Table 2 Association of multifactorial PFAS with diabetes

chemical	OR	95% CI)	P	VIF
PFDeA	0.94	(0.59, 1.47)	0.775	3.52
PFHxS	0.78	(0.68, 0.89)	0.000	1.30
MPAH	0.97	(0.75, 1.25)	0.800	1.26
PFNA	1.14	(0.91, 1.43)	0.255	2.21
PFUA	0.93	(0.58, 1.47)	0.744	2.91
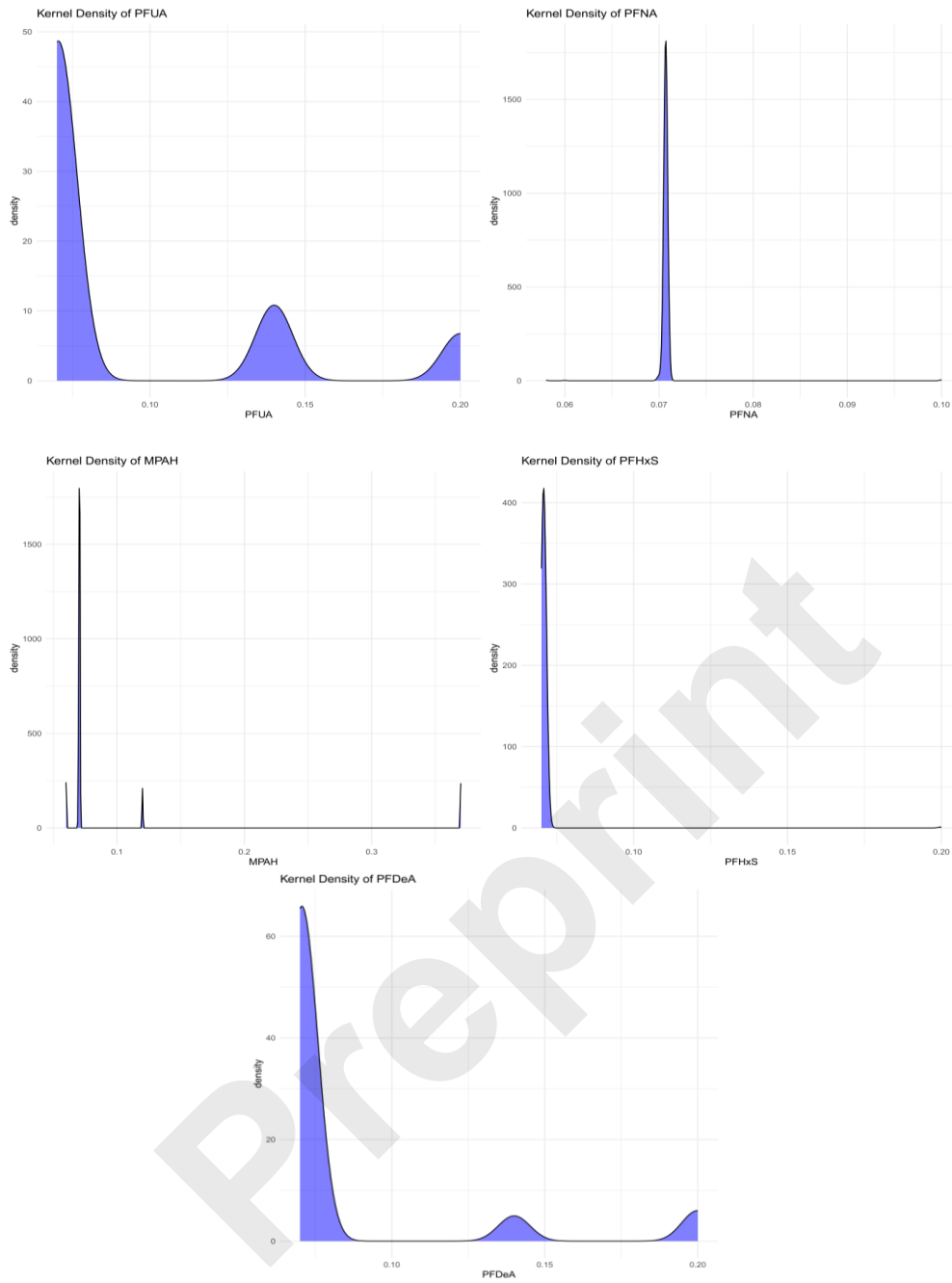
Supplementary Figure 1 Kernel function density plot of chemical measurements
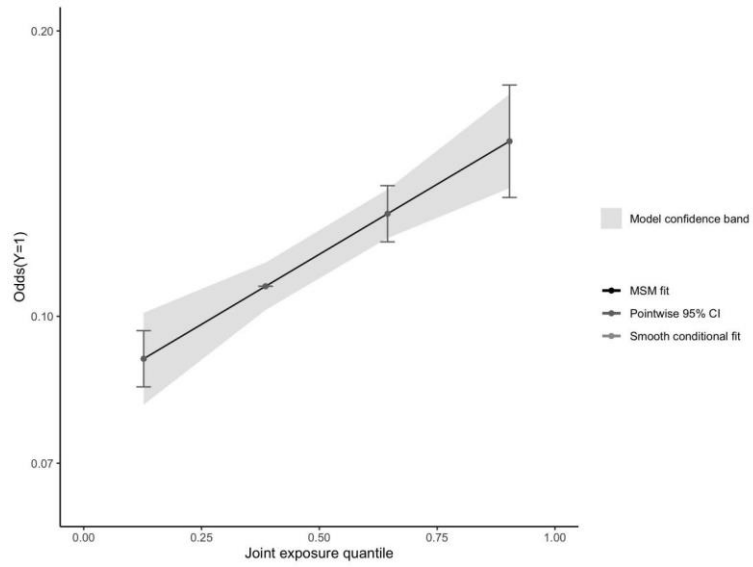
Supplementary Figure 2 Confidence bands for prevalence OR values of PFAS and diabetes
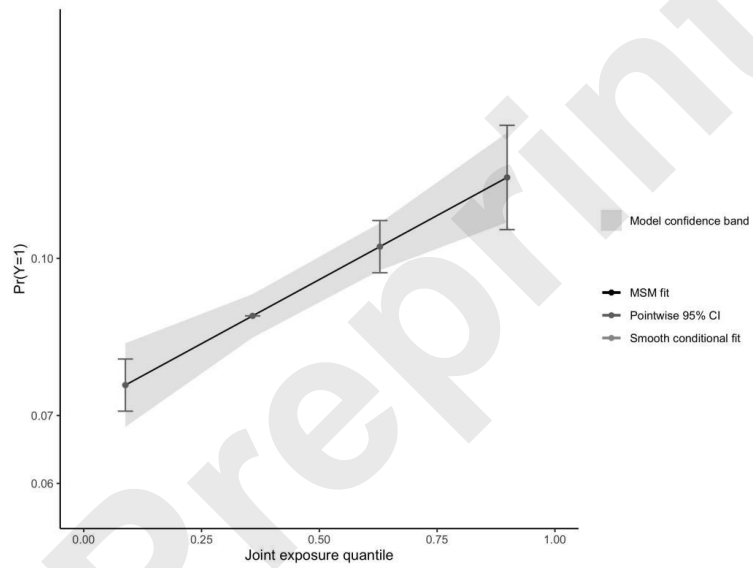
Supplementary Figure 3 Confidence bands of prevalence Pr values for PFAS and diabetes

Supplementary Table 1 Association of individual PFAS with diabetes, Model I

chemical	Ln-transform OR (95%CI)	P	Quartile1	Quartile2 OR (95%CI)	P	Quartile3 OR (95%CI)	P	Quartile4 OR (95%CI)	P
MPAH	0.66 (0.51, 0.85)	0.002	ref	0.85 (0.79, 0.91)	0.000	0.89 (0.84, 0.94)	0.000	0.76 (0.67, 0.85)	0.000
PFDeA	0.71 (0.54, 0.92)	0.011	ref	0.87 (0.81, 0.94)	0.000	0.91 (0.86, 0.96)	0.000	0.89 (0.79, 1.00)	0.047
PFHxS	0.68 (0.6, 0.77)	0.000	ref	0.78 (0.73, 0.85)	0.000	0.83 (0.78, 0.88)	0.000	0.71 (0.63, 0.8)	0.000
PFNA	0.79 (0.67, 0.93)	0.004	ref	0.91 (0.85, 0.98)	0.014	0.93 (0.88, 0.98)	0.005	0.85 (0.76, 0.96)	0.007
PFUA	0.65 (0.47, 0.89)	0.007	ref	0.83 (0.77, 0.89)	0.000	0.88 (0.84, 0.93)	0.000	0.83 (0.74, 0.94)	0.002

Supplementary Table 2 Association of individual PFAS with diabetes, Model II

chemical	Ln-transform OR (95%CI)	P	Quartile1	Quartile2 OR (95%CI)	P	Quartile3 OR (95%CI)	P	Quartile4 OR (95%CI)	P
MPAH	0.71 (0.55, 0.91)	0.008	ref	0.86 (0.8, 0.93)	0.000	0.90 (0.85, 0.95)	0.000	0.79 (0.7, 0.89)	0.000
PFDeA	0.82 (0.63, 1.06)	0.122	ref	0.93 (0.86, 1)	0.057	0.95 (0.9, 1)	0.074	0.98 (0.87, 1.1)	0.714
PFHxS	0.75 (0.66, 0.86)	0.000	ref	0.83 (0.77, 0.9)	0.000	0.87 (0.82, 0.92)	0.000	0.78 (0.68, 0.88)	0.000
PFNA	0.84 (0.71, 0.99)	0.035	ref	0.95 (0.88, 1.03)	0.199	0.95 (0.9, 1.01)	0.101	0.91 (0.8, 1.02)	0.118
PFUA	0.79 (0.59, 1.06)	0.123	ref	0.89 (0.83, 0.96)	0.003	0.94 (0.89, 1)	0.038	0.95 (0.84, 1.07)	0.391

Supplementary Table 3 Association of individual PFAS with diabetes, Model III

chemical	Ln-transform OR (95%CI)	P	Quartile1	Quartile2 OR (95%CI)	P	Quartile3 OR (95%CI)	P	Quartile4 OR (95%CI)	P
MPAH	0.92 (0.71, 1.19)	0.528	ref	0.98 (0.89, 1.07)	0.653	0.97 (0.91, 1.03)	0.320	0.94 (0.81, 1.08)	0.395
PFDeA	0.92 (0.71, 1.18)	0.499	ref	0.99 (0.92, 1.07)	0.810	0.99 (0.94, 1.06)	0.934	1.08 (0.95, 1.22)	0.257
PFHxS	0.79 (0.7, 0.9)	0.000	ref	0.86 (0.79, 0.93)	0.000	0.89 (0.84, 0.95)	0.000	0.81 (0.72, 0.92)	0.201
PFNA	0.96 (0.82, 1.13)	0.646	ref	1.02 (0.94, 1.11)	0.574	1.01 (0.95, 1.07)	0.792	1.01 (0.89, 1.15)	0.821
PFUA	0.92 (0.69, 1.21)	0.537	ref	1.02 (0.93, 1.11)	0.734	1.00 (0.95, 1.06)	0.965	1.07 (0.94, 1.22)	0.285

Supplementary Table 4 Association of WQS positive mixture coefficients of PFAS with diabetes

model	Estimate	Std. Error	z	$P\,(>|z|)$
no ajusted	1.02	0.04	0.46	0.643
model1	1.77	0.04	-5.79	0.000
model2	1.83	0.04	-4.18	0.000
model3	1.83	0.04	-4.11	0.000

Supplementary Table 5 Association of WQS negative mixture coefficients for PFAS with diabetes

model	Estimate	Std. Error	z	$P\,(>|z|)$
no ajusted	1.99	0.04	-0.34	0.737
model1	1.87	0.04	-3.70	0.000
model2	1.91	0.05	-2.07	0.000
model3	1.90	0.05	-2.13	0.000

Supplementary Table 6 Different models of PFAS Quantile g-computation OR for mixture effects

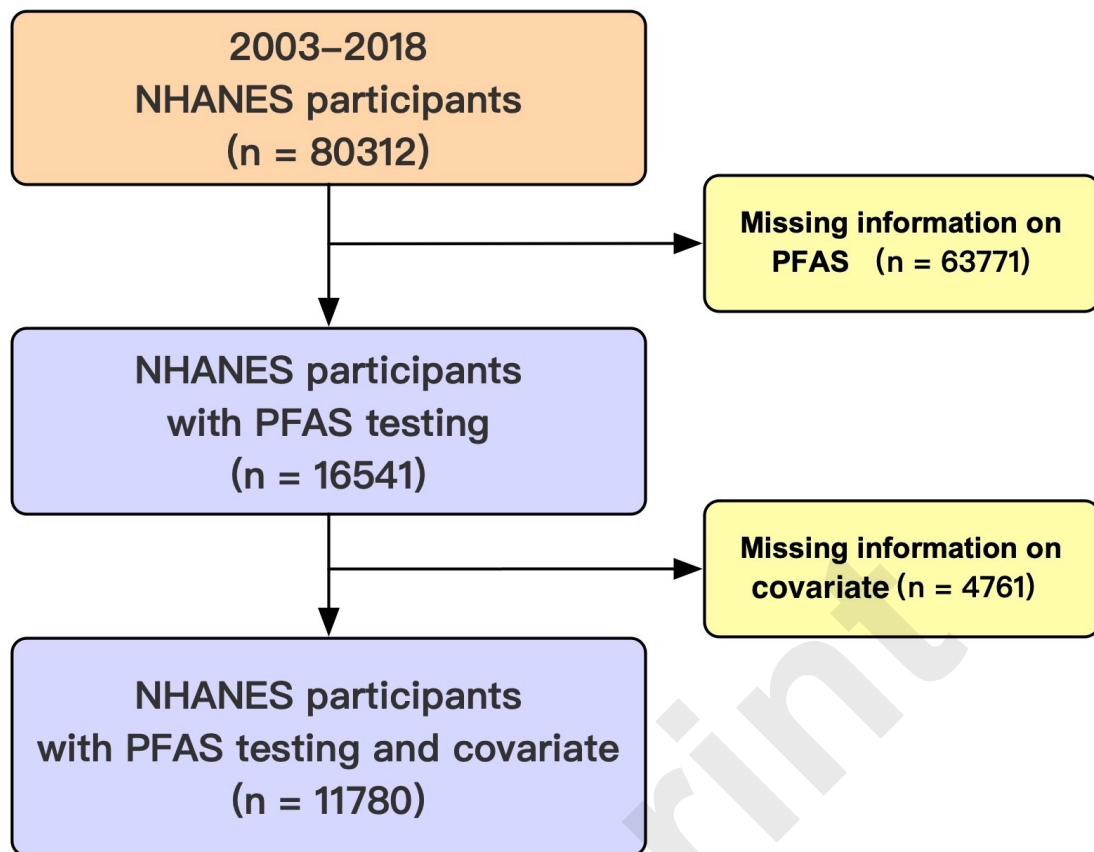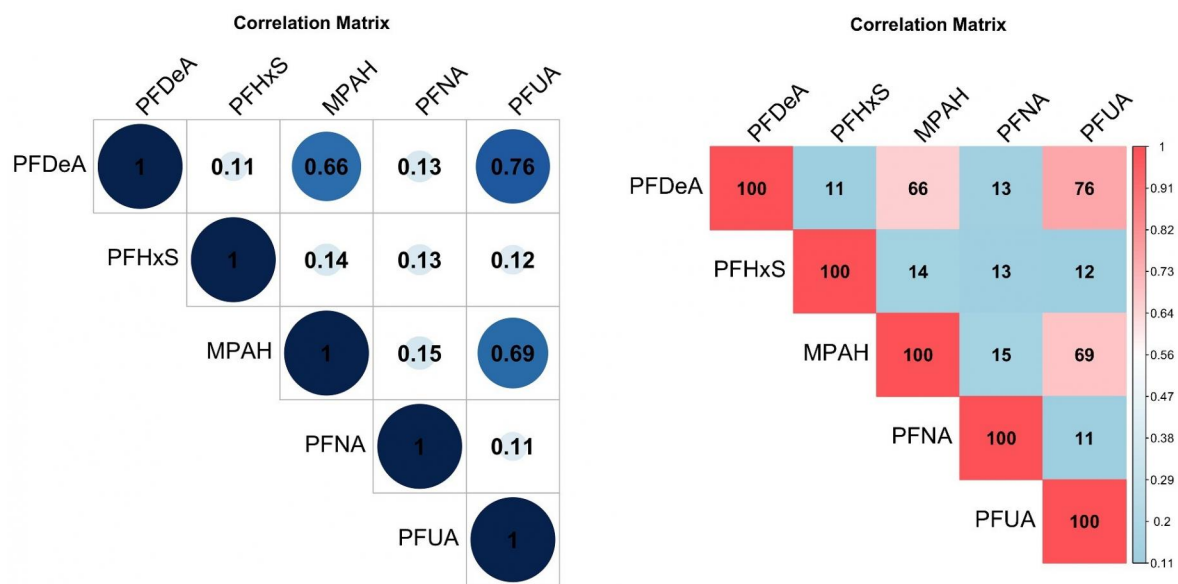|	coefficients. Estimate	coefficients. Std. Error	coefficients.Lower.CI	coefficients.Upper.CI	coefficients. Z. value	coefficients. Pr. z
model 1	1.78	0.04	1.72	1.84	-6.70	2.04E-11

Fig. 1 Inclusion and exclusion path diagram

Fig. 2 Spearman correlation coefficient between pairwise chemical measurements
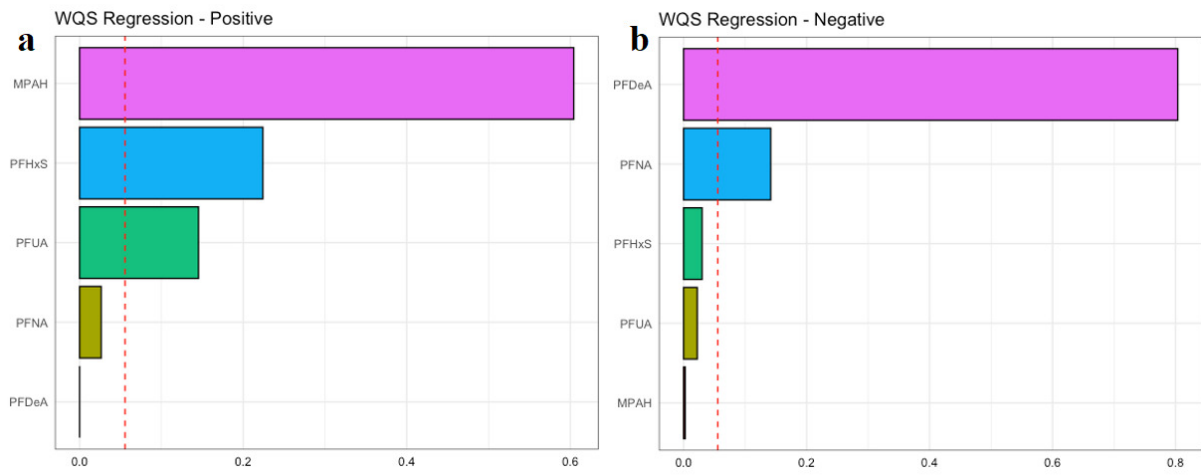
Fig. 3 Weight of regression indicators on diabetes in the WQS model. (a) Forward constraints; (b) Negative constraints
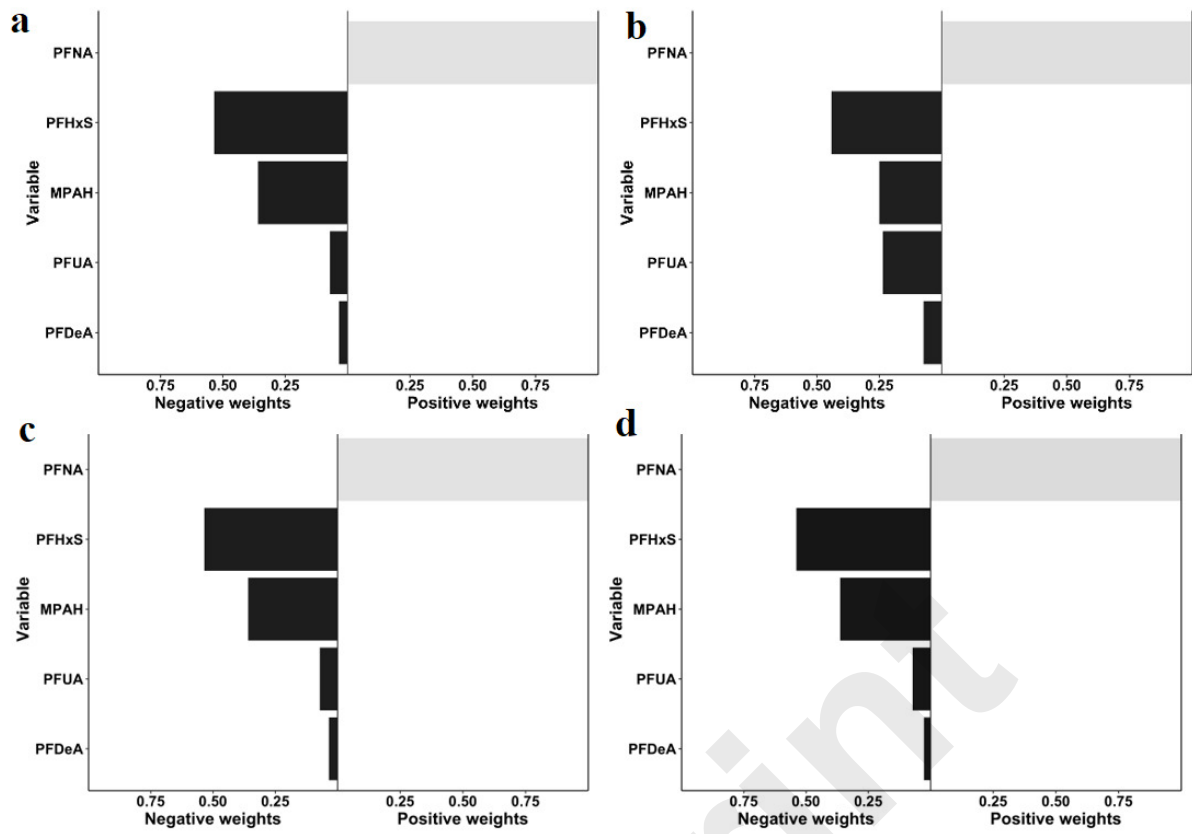
Fig. 4 Q G-computation weights for the four models. (a) Model I; (b) Model II; (c) Model III; (d) Model 4